(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US];

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP) 3 February 2000 (03.02.2000) Filed on US 09/560,875 (CIP)

Filed on

27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom -{US/US}; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, butilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

5

10

15

20

25

30

35

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that. have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

10

15

20

25

30

35

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

5

10

15

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5

10

15

20

25

30

35

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

25 4.3 ANTISENSE

5

10

15

20

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

5

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

5

10

. 15

20

25

30

35

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

5

10

15

20

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

5

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

35

10

5

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP 20 (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. 25 Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available 30 from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

5

10

15

20

25

30

35

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

5

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

- J,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

10

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No.

WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

30

35

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

10

15

20

25

30

35

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeck, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

5

10

15

20

25

30

35

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

5

10

15

20

25

30

35

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (*i.e.*, increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

5

10

15

20

25

30

35

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- 10 (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

15

20

25

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particularneurotoxins; and
 - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

- (ii) increased sprouting of neurons in culture or in vivo:
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape

35 (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15 4.10.19 IDENTIFICATION OF POLYMORPHISMS

5

10

20

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

10 4.10.20 ARTHRITIS AND INFLAMMATION

5

15

20

25

30

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

10

15

20

25

30

35

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

5

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

10

15

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

5

10

15

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation, See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, 35 synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol., 133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

5

10

15

20

25

30

35

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, <u>J. Mol. Biol.</u>, <u>227</u>:381 (1991); Marks et al., <u>J. Mol. Biol.</u>, <u>222</u>:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (<u>Bio/Technology 10</u>, 779-783 (1992)); Lonberg et al. (<u>Nature 368</u> 856-859 (1994)); Morrison (<u>Nature 368</u>, 812-13 (1994)); Fishwild et al.(<u>Nature Biotechnology 14</u>, 845-51 (1996)); Neuberger (<u>Nature Biotechnology 14</u>, 826 (1996)); and Lonberg and Huszar (<u>Intern. Rev. Immunol. 13</u> 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)/2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)/2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

5

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

5

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
 - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

5

10

15

20

25

30

35

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, *e.g.*, Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategics. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ($CviII^{**}$), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a $CviII^{**}$ digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that $CviII^{**}$ restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

· 10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers *e.g.* a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

20 5.0 EXAMPLES

5

10

15

25

30

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library N	lame SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
			976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
			217 225 238 271 317 404 446 469 503
			513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
adult brain	Clontech	ABR001	1254 39 216 238 327 356 535 927 1056 1121
addit Oralli	Clonicen	ABROOT	1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
		•	147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
	[•	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
		Ì	658 666 683 725 742 764 767 786 801 805 810 823 826 829 831 836 841 887-
		İ	923 927 934 943 950-951 963 976 995
	1		1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
			1128 1142 1162 1181-1192 1199 1204
	Ì	ł	1218-1219 1225 1232 1253 1267 1271-
			1306 1342 1347 1349-1350
adult brain adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain Invitrogen	ABR012 ABR013	74 238
adult brain	Invitrogen	ABT004	868 1268 49 117 138 191 217 252 291 305 535
adult brain	nividogen	AB1004	566 596 663 670 746 798 816-819 876
	j		892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes		İ	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
			240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
	OBSS	THE COT	118 129 132 138 151 158-163 182 195-
	İ		203 215 217 238 264 269 353 384 398
	Ì	Ì	408 434-439 446 504 512-513 519 537
		ł	562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
	i	}	874 891 898 919 926-927 976 988
			1021 1037 1041 1062 1067 1071 1080 1083 1093 1122 1131 1185 1201 1254
			1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
			107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
			446 454 477 504-505 509 514 518-519
	Ţ		535 537 564 574-583 620-627 639 653
			673-675 705 753 789 831 844 851 859
		1	877 909 918 927 956 963 976 1067
adult kidney	Invitrogen	AKT002	1074 1083 1095 1178 1302 1331 1335
addit vidites	mainogen	AK 1002	446 487 564 575 844 868 910 927 976
			1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514
			1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

Tissue Origin	RNA Source	Hyseq Library Na	me SEQ ID NOS:
		 	518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
_			545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
			519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
	1		505 657 675 714 753 832 844 941-942
			976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
	1		104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264
		į (316 384 409 440 445-446 496 504 512
			514 518-519 535 537 549-550 564 566
			571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771
			815 833 842-865 868 875-876 918 926-
	ľ		927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
			1124 1131 1144 1174 1224 1268 1331
<u></u>			1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
			294 414 446 477 504 514 534 545 549
			592 722 873 883 952 976 1041-1042
 		1,00000	1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
			238 446 497 537 642 701-706 811 877
		DI DOO!	927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
,	Clontech	DIADON	592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
	ĺ		147 151 164-174 213-215 238 305-307 374 404 446 460 466 516 519 534 538-
	-		
	· }		541 544-546 549-554 566 584 586 592 596 607 610 628-629 643-645 652 707-
			708 774-789 844 866-871 873 919 927
	1		952 963 976 998 1034 1042 1064 1083
			1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
bone manow	Cioniccii	151415002	210 317 510-511 545 549 581 598 628
		}	638 724 766 789 844 860 868 873 919
	İ	•	927 952 963 968 976 1042 1111 1141
	j]	1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	111 238 282 549 1083
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
addit cololi	HIVIOGOI	CLITOUI	844 873 877 952 976 1042 1152 1268
	1	1	1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
COUNT CEL AIX	Diocham	CYNOVI	335 365-367 392-401 440 466 470-471
	1		518 537 597 629 832 877 927 976 1006
			1085 1117 1129-1134 1192 1202-1205
		1	1219 1309-1328
diaphragm	BioChain	DIA002	74 976 1083
	i Diochail	1 1/1/1/1/2	1 1 7 7 10 1003

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
		· · -	138 151 204-206 215-217 238 269 316
1	1	}	414 433 505 510 513 550 555 580 582
		1	596 675 722 745 798 814 836-841 851
ľ	1	{	918 976 1041 1043 1073 1083 1131
	1		1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic	<u> </u>	1
of chromosome 8 Genomic clones	Research Genomic DNA	EPM003	47 525
from the short arm	from Genetic	LI MIOOS	77 323
of chromosome 8	Research		1
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research	!	
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	1	(
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
	1	l	225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
	1		829 836 859 909 927 943 947 963 1057 1067-1068 1104 1135-1140 1162 1206-
	}		1067-1068 1104 1135-1140 1162 1206-
		ļ	1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
_		- · -	535 683 761 798 820-827 844 876 909
	1	ŀ	963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844 859 1048 1083 1116 1193
fetal liver-spleen	Columbia	FLS001	859 1048 1083 1116 1192 8 14 34-35 37 41 43 49 51 54-56 63-64
Town HACL-Shieeli	University	1.50001	69-71 74 77 79 87-90 101 107 110-111
	Chrossity		114 120 128-131 138 140 147 150-155
1	 		197 210 215 217 225 238 312 367 384
	1	Ì	414 440 446 460 468 483 496 504-507
[[1	511-515 518-519 523 533-535 537 541
i	1	ļ	544-545 547-550 555-560 564 566 571
i	1		577 582 585-586 598 636 646-647 649
ı	1	' ·	652 664 698 709-710 714 722-723 731
ì	}		735-736 746-753 761 784 798 823 829
ŀ	ļ .		832 844 851 858-859 868 873 876 898
1	į l	1	927 943 949 952 963 976 984 1002
ŀ		•	1021 1023 1040 1042 1044 1050 1083
į			1093 1116 1120 1129 1131 1144 1174
i	, 1	}	1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
	University	- 20002	111 120 129 147 207 210 215-216 238
i			250 330 353 359 366 383-384 414 478
į l	1		505 508-509 511 515-524 534-535 537
			544-545 564 566 571 577 591 598 638

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
	12373000	-23004 Esoluty Hullio	663 671 698 714 722 725 727 751 798
	l		851 859 873 876 909 927 949 952 983-
	1		984 1002 1023 1042-1044 1085 1095
	1		1131 1144 1178 1199 1233 1240-1270
			1331 1340
fetal liver-spleen	Columbia	FLS003	64 535 976 1256
	University		
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
			580 722 730 749 844 918 943 976 1051
			1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141
			1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
			225 264 316 405 422-429 488-494 496
İ	1		519 534-535 537 566 675 732 859 876-
			877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
	1		1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
	}		316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
			976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
	ł		138 145 151 188 197 207 215 238 264
	1	1	271 294 316 367 414 440 446 466 504
	ļ		513-514 535 542-543 550 564 571 596
	j		635 648-654 675 711-715 722-723 798
	Į		832 872 876 883 927 976 1095 1144
_		<u> </u>	1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
	}		299 308-310 371-373 462 476 504 511-
			513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
)	858-859 876 898 909 949 976 1045-
			1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
· · · · · · · · · · · · · · · · · · ·			1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
	<u> </u>	<u> </u>	1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
	University		
infant brain	Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
	University		1221
lung , fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
	1		215 217 269 280 296 337 374-375 384
			404 446 454 475-480 498 514 518-519
			522 537 545 564 577 597 653 658 705
	1		721-724 754-756 779 859 868 872-874
	1		876-877 919 927 949 951-952 959 976
	}		1002 1042 1048-1053 1076 1083 1088-
		Į.	1089 1131 1144-1147 1216-1218 1229

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		T	1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
		1	634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
•			147 151 212 215 218 238 252 288 312-
	•	}	314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
	}		564 566 571 577 580 582 587-609 615
	1		632-638 658-659 698 714 725-728 832
	•		836 841 859 866 873-874 882-883 918-
	1		919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
-	1		657 660 729 873 883 927 952 963 1008
	1		1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	1		919 929 939 952 976 1071 1118 1218
1424	1		1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
, , , , , , , , , , , , , , , , , , , ,			217 250-256 264 297-299 305 377-378
	}		398 446 481-486 505 512 537 545 549
	1	•	571 592 725 730-733 816 829 836 844
	ł		868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
	}		1055 1076 1083 1091 1093 1116-1117
			1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
			1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells			}
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
•	·		1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
		1	1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
, 0	1		976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
			963 967-978 1042 1120 1152 1223-
			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
-			270 343-344 353 379 516 537 566 740
	1		828 927 976 979-994 1092 1153-1159
	ļ		1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
			995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
			963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
			750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
)-1140	0101110011		317-319 336 340 359 380 398 446 448-
	}	1	463 512 519 545 554 587 598 698 724-
: (l		725 789 812 836 868 873 927 947 952
	L		125 ,07 012 030 000 013 721 741 732

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
		1	1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
	j		210 217 222 253 264 271 277-286 294
			320-326 345-352 361 381-382 446 467
ĺ			483 514 534 549-550 564 578 602 649
			844 882-883 927 950 956 976 1008-
ļ.	-	1	1028 1076 1083 1117-1120 1142 1163-
		<u> </u>	1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
		1	952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
	1		885-886 976 1001 1032-1033
			1232

TABLE 2

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:		 		Score	
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ	Accession	Species	Description	Smith-	1%
ID NO:	No.		•	Waterman Score	Identity
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
31	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	95
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus annuus	hydroxyproline-rich protein	110	31
45	U82288	Caenorhabditi s elegans	Rac-like GTPase	139	70
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63 64	U35376 AF265555	Homo sapiens Homo sapiens	repressor transcriptional factor ubiquitin-conjugating BIR-domain enzyme	485 785	65 74
45	CORRER	170	APOLLON	00	05
66	G03883 AF177390	Homo sapiens Manduca sexta	Human secreted protein, SEQ ID NO: 7964. antennal specific membrane protein AMP	274	95 54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70	W75770	Homo sapiens	Human oxidoreductase YTFO3.	1144	98
71	AB011135	Homo sapiens	KIAA0563 protein	239	76
72	AB014885	Halocynthia roretzi	HrPOPK-1	813	78
73	AF045454	Cavia porcellus	phospholipase B	955	73
74	J02870	Mus	laminin receptor	308	61

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	<u> </u>	<u> </u>		Score	
	ļ	musculus		l	
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
-	15110054	norvegicus	<u> </u>		
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
	Y38422	TTo me appliance	complex component TRAP240	468	-
77	Y14596	Homo sapiens	Human secreted protein. Human T-type voltage-gated Ca channel alpha-	1357	76 99
78	1 14396	Homo sapiens	1-I (hCavT3).	1337	99
79	Y14591	Human	APM-1 protein	767	100
		papillomaviru	-	į	
		s type 68)	
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
	<u> </u>	thaliana			<u></u>
82	L46815	Mus	DNA binding protein Rc	895	75
		musculus			<u> </u>
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
05	AB029002	17	designated HSCOP-6.	122	142
85 86		Homo sapiens	KIAA1079 protein	134 325	62
86	Y28678 Y99368	Homo sapiens	Human cw272_7 secreted protein. Human PRO1326 (UNQ686) amino acid	156	48
01	1 3000	Homo sapiens	sequence SEQ ID NO:100.	130	40
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
00	AJ223124	musculus	HAC3	107	1 33
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
•	1	brucei			[
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
		norvegicus		l	
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	675	48
100	1 500000	norvegicus	kinase	867	ļ
100	AF279265 AC007878	Homo sapiens	putative anion transporter I match to nuclear protein, NP220; note: sequence	160	98 60
101	AC007878	Homo sapiens	difference at residue 58	100	00
102	U22829	Mus	P2Y purinoceptor	264	42
102	1 22027	musculus	2. parmocopioi	~~	""
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
			receptor-B3.	1	
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
			119 SEQ ID NO:119.	 	
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia	sialic acid transporter	587	96
		coli			
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	182	94
III	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
112	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein	274	51
	1,775	Tromo suprous	sequence SEQ ID NO:84.	1	1
113	AF016365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AC007956	Homo sapiens	unknown	520	75
115	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets-	484	91
				1	1
	Į.		domain transcription factor ESE-3A, isoform 1))	i	

SEQ	Accession	Species	Description	Smith-	1%
m `	No.			Waterman	Identity
NO:])	Score	100,111,
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
	1.2000/10	norvegicus	prospinancy into site 5 km and 6	021	1
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr	1646	94
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi s elegans	contains similarity to C2 domains	219	29
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IF116b	496	67
131	AF201734	Mus musculus	testis specific serine kinase-3	800	87
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133 -	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1157	87
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860)	5041	99
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDII-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	392	61
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
160	AJ250425	Rattus norvegicus	Collybistin I	556	74
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100

SEQ	Accession	Species	Description	Smith-	T%
ID	No.	- Prince		Waterman	Identity
NO:	1	1		Score	
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus	calcium transporter CaT1	700	96
166	Y76332	norvegicus Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus	NIK-related kinase	197	43
169	AF252293	musculus Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
· · _ · _		gen. sp.			
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55 -
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
179	AF127481	Homo sapiens	codon) non-ocogenic Rho GTPase-specific GTP	175	80
		l	exchange factor		
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone hi968 2.	382	49
196	AF255614	Rattus	scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
198	AF073967	Mus musculus	olfactory receptor	316	43
105	1	domesticus	-	 	
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
205	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.		-	Waterman	Identity
NO:	<u> </u>			Score	
		ļ <u>-</u>	{ovarian cancer critical region of deletion}		
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific protein, calphotin.	541	82
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
	1 2200	norvegicus	neurofascin	700	1
213	AF154846	Homo sapiens	zinc finger protein	798	56 93
214	AF102777	Mus musculus	FYVE finger-containing phosphoinositide kinase	933	
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	563	78
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus musculus	Kupffer cell receptor	567	40
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	636	96
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus	•		ļ
226	AE000218	Escherichia coli	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus musculus	GTP-binding like protein 2	265	88
229	AF122924	Xenopus laevis	Wnt inhibitory factor-1	316	40
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific phospholipase-D.	290	100
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus musculus	Embryo Brain Kinase	460	62
238	U64857	Caenorhabditi s elegans	similar to the BPTI/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor (TFPI)	284	33
239	AJ250840	Mus musculus	serine/threonine protein kinase	739	63
240	AJ223472	Mus musculus	transcription elongation factor TFIIS.h	222	38
243	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein sequence SEQ ID NO:18.	353	52
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
- · -		norvegicus		1	1
244	AF016191	Rattus norvegicus	potassium channel	1043	98
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	645	98
246	Y29868	Homo sapiens	Human secreted protein clone pp325 9.	497	98
247	AF180475	Homo sapiens	Not4-Np	188	83
248	Y17227	Homo sapiens	Human secreted protein (clone yal-1).	690	99
249	AF250910	Manduca	death-associated small cytoplasmic leucine-rich	182	31

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	ļ		- COLD	Score	
250	AF192756	sexta	protein SCLP	124	34
250	AF 192/56	Kaposi's sarcoma-	Orf73	134	34
		associated		1	1
		herpesvirus			
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus	DNA binding protein Rc	251	67
233	1,40813	musculus	DIVA binding protein KC	231	0/
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus Mus	Citron-K kinase	1201	98
233	AI-070000	musculus	Cit on-K kindse	1201	76
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus	Phospholipase	368	80
23,	2,2041	cuniculus	i nospitenpase	1 300	00
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus	L-periaxin	430	72
	1.4222700	musculus	· ·	1.50	'-
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
	1	norvegicus		1	1
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
	1	110	(GPCR).	1	
266	U27269	Mus	sodium glucose cotransporter	204	56
		musculus		1	
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus	putative taste receptor TR1	209	39
	}	norvegicus	•	1	j
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
	l	pyogenes	·	<u> </u>	
271	AB009883	Nicotiana	KED	109	26
	<u> </u>	tabacum			
272	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
	1	musculus		<u> </u>	
273	L34938	Rattus	ionotropic glutamate receptor	460	86
-	1	norvegicus		L	<u> </u>
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent	188	74
	1		Expressed Protein LIKE PUTATIVE protein)	ì	1
275	AF265555	17		1	94
213	Ar203333	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	173	94
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens		283	96
279	AC008075	Arabidopsis	KIAA1631 protein Contains PF 00069 Eukaryotic protein kinase	157	43
417	1 2000013	thaliana	domain.	131	43
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	439	84
283	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
200	72 130330	musculus	Dro-goniam nanscribnona rebiessor LE1	1 303	"
284	Y29336	Homo sapiens	Human secreted protein clone cs756 2 alternate	647	100
~177	127330	Tronto sapiens	reading frame protein.	1 34,	100
	Y73402	Homo sapiens	Human secreted protein clone ye25_1 protein	300	90
285	113402	Tollo subicits	sequence SEQ ID NO:26.	300	1 70
285	i			1	L
	AF016411	Homo coniero	KCNA3 1B	137	1 100
286	AF016411	Homo sapiens	KCNA3.1B	137	100
286 287	W89253	Homo sapiens	Human ALP.	688	97
286 287 288	W89253 AF112886	Homo sapiens Bos taurus	Human ALP. differentiation enhancing factor 1	688 750	97 96
286 287	W89253	Homo sapiens	Human ALP.	688	97

SEO	Accession	Species	Description	Smith-	1 %
SEQ		opecies	Description		
ID	No.			Waterman	Identity
NO:		<u> </u>		Score	<u> </u>
		norvegicus			ļ
292	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
	ļ	norvegicus	interacting protein 2 Maguin-2		<u> </u>
293	X99211	Drosophila	ubiquitin-specific protease	143	38
		melanogaster			ł
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	185	94
	1		sequence SEQ ID NO:92.		
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577_1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14 1 protein	182	97
	1		sequence SEQ ID NO:92.		
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	605	69
200	100000	110ino sapions	gene 16 SEQ ID NO:63.	1 003	0
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859		neuropilin-2(a0)	277	100
		Homo sapiens			
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
	i		Method: conceptual translation supplied by		İ
	i		author		
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
	ľ	musculus		ţ	
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
	i	norvegicus			
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
J	005.75	discoideum	outside protein		50
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
J12	10/3//	Tromo suproms	124 SEQ ID NO:124.	' ' '	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
314	ACCOUNT	Tionic sapiciis	44% similarity to U42767 (PID:g1736918)	'''	1 30
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and	278	38
313	AL021392	Homo sapiens	GENEWISE)	270	1 30
316	U70209	Mus	polycystic kidney disease 1 protein	165	38
310	070209		polycystic kioney disease i protein	165	36
217	A E 100642	musculus		202	30
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
	1	norvegicus			<u> </u>
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma	activated protein kinase C receptor homolog	141	38
		vivax			
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
	i		receptor		1
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
		•	123.]	
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214	97
		Tromo suprens	3.1.4.37)	l :	´´
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756		protein kinase C mu	540	78
		Homo sapiens			
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable	581	80
	L		region		
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP-	1127	100
	<u> </u>	_	107 SEQ ID NO:107.	<u>_</u>	l
333	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	320	98
334	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	327	93

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
			similarity to P49205 (PID:g1345860)		
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	1111	67
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus musculus	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+- activated potassium channel	463	80
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361 362	R28916 U16655	Homo sapiens Rattus	Type III procollagen (prior art). phospholipase C delta-4	108 649	40 65
363	G03119	norvegicus Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	1		Waterman	Identity
NO:	1	'		Score	
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus	GTP binding protein	1456	91
370	7132374	musculus	OII omang protein	1430	1 ''
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830		protein tyrosine kinase	115	44
382	U0483U	Dictyostelium discoideum	protein tyrosine kinase	113	44
202	G02916		11	(10	98
383		Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by	171	34
	ĺ	1	gene 38.		
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates	dopamine receptor D4	105	35
-,.	1.2002,00	syndactylus	- Copenial Cooper	1 .02	1 55
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by	1047	92
577	151.05	i iomo supiems	gene 2 SEQ ID NO:126.	1 .0.7	1 12
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein;	527	78
401	D87002	Tionio Sapiciis	accession number 221513.	321	/*
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus	ADP-ribosylation factor-directed GTPase	545	89
404	A1073402	musculus	activating protein isoform b	343	07
405	X92887	Human	pol/env	162	30
403	A92007	endogenous	povenv	102	30
	ļ	retrovirus K		ł	Į.
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
400	AK022626	Homo sapiens	unnamed protein product	2833	99
407				4	
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by	1788	89
410	11/00245	77	gene 9 SEQ ID NO:273.	2004	-
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone	2004	99
433	1 70 10050	 	HTSEV09.		ļ
411	AB043953	Mus	Chat-H	2628	82
		musculus		ļ. <u></u>	ļ
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID	1014	92
	L		NO:148.		
413	U10542	Pan	MHC class I A	265	71
		troglodytes			
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia	alpha tubulin	289	68
		coriiceps			
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar	1446	96
421	1 107051				I .
421	ALTOSOZI	1	to rat sperm antigen 4 (SPAG4)))		1
421 422	AC008075	Arabidopsis	to rat sperm antigen 4 (SPAG4))) F24J5.4	112	35

SEO	Accession	Species	Description	Smith-	1%
ID	No.	Брсска	Description	Waterman	Identity
NO:	140.			Score	lucinity
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ	1961	99
723	133742	Tionto sapiens	ID NO. 191.	1501	1 "
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
430	104441	Fiolilo Sapiciis	associated protein.	2074	100
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon	extensin-like protein	613	48
433	AF139290	esculentum	extensin-like protein	1 013	40
434	W48351	Homo sapiens	Human broad cancer related protein PCPP2	135	44
435	X73874		Human breast cancer related protein BCRB2.	3442	97
		Homo sapiens	phosphorylase kinase		
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus	transmembrane receptor UNC5H1	817	93
		norvegicus			ĺ
450	AF081249	Homo sapiens	JAW1-related protein MRVIIA long isoform	4568	99
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
	İ	1	(CIRP-1).		
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
	1	J	gene 62.	j] _
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum		ļ	ł
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
		<u> </u>	clone HTDAD22.		<u></u>
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
		<u> </u>	gene 17.	<u> </u>	
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein	135	47
		1	sequence SEQ ID NO:16.		
463	X84960	Triticum	low molecular weight glutenin	109	33
		aestivum	<u> </u>	_	
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus	alpha/beta hydrolase-1	502	59
	<u></u>	musculus]
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	99
	1	1	gene 77.		l
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus	neurotoxin homologue	118	47
	1	musculus			
			•		
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1	į	Waterman	Identity
NO:				Score	<u> </u>
473	G02313	Homo sapiens	gene 62.	220	1.00
474	Y07007	Homo sapiens	Human secreted protein, SEQ ID NO: 6394. Breast cancer associated antigen precursor	328 1013	97
4/4	10,007	Homo sapiens	sequence.	1013	197
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
	1		clone HTDAD22.	202	100
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
	<u> </u>	musculus	<u> </u>	L	
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481.	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
		<u> </u>	SYTAX1.	<u> </u>	<u> </u>
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483 484	AF210651	Homo sapiens	NAG18	124	59
485	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485 486·	G00637 U15174	Homo sapiens Homo sapiens	Human secreted protein, SEQ 1D NO: 4718. BCL2/adenovirus E1B 19kD-interacting protein	129	70
400	013174	rionio sapiens	BCL2/adenovirus E1B 19kD-interacting protein 3	149	73
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	197	67
495	AC005175	Homo sapiens	R31449_3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis familiaris	D4 dopamine receptor	90	48
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus maniculatus	reverse transcriptase	213	52
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135_9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	115	33
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light polypeptide kinase))	184	92
509	U43360	Peromyscus maniculatus	reverse transcriptase	97	62
510	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
511	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
513	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila melanogaster	CG6393 gene product	259 .	42 .
515	Z17206	Xenopus laevis	p46XIEg22	128	40
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
518	AF151083	Homo sapiens	HSPC249	444	98
519	S80864	Homo sapiens	cytochrome c-like polypeptide	318	50
520	X92485	Plasmodium	pval	170	61
	1	vivax	•		l

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	operios	Jaconpuon	Waterman	Identity
NO:	1	1		Score) lucinity
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857	Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	253	73
		110	HPMBQ32.		1.
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528	U47924	Homo sapiens	C8	1112	86
529	G04063	Homo sapiens	Human secreted protein, SEO ID NO: 8144.	84	45
530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	gin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi	contains similarity to a BR-C/TTK domain	853	39
5.0	1.010.50	s elegans	Contains similarly to a Dix G 1 112 donain	033	1 "
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN: 45%	408	66
J.,	110005075	Tromo saprons	similarity to P22059 (PID:g129308)	100	1 55
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus	olfactory receptor F3	327	73
	}	musculus			''
544	Y73431	Homo sapiens	Human secreted protein clone yb186 1 protein	386	100
	1		sequence SEQ ID NO:84.		1
545	AE004833	Pseudomonas	probable TonB-dependent receptor	279	42
	1	aeruginosa	•	1	1
546	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein	1772	67
		•	B receptor protein.	ļ	1
548	¥91493	Homo sapiens	Human secreted protein sequence encoded by	176	100
		•	gene 43 SEQ ID NO:166.	ſ	-
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584 2 protein	1224	94
	į.	· ·	sequence.		
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
554	AB025258	Mus	granuphilin-a	501	41
	İ	musculus	ļ.	j	
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein	183	32
	1	1	complex component TRAP150	1	}
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus	Ca2÷ dependent activator protein for secretion	1010	93
_	1	musculus		[İ
561	AF187325	Canis	melanoma antigen	287	55
	1	familiaris	_	Į	}
562	AJ001981	Homo sapiens	OXAIL	2512	99
563	Z17238	Rattus	glutamate receptor subtype delta-1	338	66
	1	norvegicus	-		
564	W30638	Homo sapiens	Partial human 7-transmembrane receptor	371	100
	1	'	HAPO167 protein.	1	1
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid	1138	78
	1	1	sequence SEQ ID NO:63.	1	1
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
307					

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:		}		Score	
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
3/1	107030	Tionio sapiciis	sequence.	1004	100
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
			diadyin 13.3 (novel protein)		
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
[ID NO:388.		
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
دەد	AB030234	familiaris	D4 dopamine receptor	04	1 20
504	C02(21		VI	245	1-00
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
		<u></u>	Antigen)		
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
		musculus	·		ĺ
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	81
		•	HPMBQ32.		
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
			protein.		
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
1 572	133073	110mo suprems	polypeptide.	130)	1 72
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	97
393	1 23031	Tionio sapiens	sequence SEQ ID NO:108.	1112	71
594	Y27658	11		763	79
		Homo sapiens	Human secreted protein encoded by gene No. 92.		58
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	1 2 7 .
596	AF151110	Mus	COPI protein	2215	95
		musculus			
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus	putative secreted protein ZSIG37	143	40
	<u> </u>	musculus			l
599	AF119855	Homo sapiens	PRO1847	236	76
600·	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
	- 22-2-5020]	Antigen)		
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
					<u> </u>
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
		<u> </u>	НРМВQ32.		L
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
			107.		L
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
			clone HTDAD22.]
614	M87053	Rattus	lens membrane protein	450	84
		norvegicus			1 -
615	AC004232	Homo sapiens	FPM315	163	37
616	G01984			205	79
010	JV1704	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	203	17

NO: 617 618 619 620 621	Y91524 AJ245621	Homo sapiens	Human secreted protein sequence encoded by	Score	
619 620 621		1	gene 74 SEQ ID NO:197.	821	99
620 621 622		Homo sapiens	CTL2 protein	2258	99
621 622	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
622	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
	D90721	Escherichia coli	Transmembrane protein dppC	573	90
	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
651	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
658	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134	65
659	Y91423	Homo sapiens	Human secreted protein sequence encoded by	333	96

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman Score	Identity
NO: 661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68
662	Y53886		A suppressor of cytokine signalling protein	375	43
002	133880	Homo sapiens	designated HSCOP-6.	3/3	43
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like) protein (hs6M1-21))		
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus	contains transmembrane (TM) region and ATP	757	68
	<u> </u>	musculus	binding region		
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KJAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	Score 125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
707	103320	Tionio sapidis	ligand (clone 2DD).	121	13
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	59
	<u> </u>	es cerevisiae			
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
712	D21211	norvegicus Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota	olfactory receptor	615	83
,,,,	1.00000	marmota	onablery receptor	1 015	1 03
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
720	AB020598	Homo sapiens	receptor beta4 subunit peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
, , ,	155000		designated HSCOP-6.	1 3/0	'-
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter;	111	41
		tigrinum	NBC		
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
70.5	V54672	musculus	protein 3A	l-3334	-
725 726	X54673 AF016191	Homo sapiens Rattus	GABA transporter potassium channel	3114 370	99 100
720	Aroloisi	norvegicus	potassium channer	370	100
727	ΔB029559	Rattus	BATI	139	35
		norvegicus		}	
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila)	142	68
731	Z10062	Homo sapiens	homolog) cDNA encoding a human vanilloid receptor	675	99
/31	210002	nomo sapiens	homologue Vanilrep1.	073	99
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia	putative transport protein	592	97
		coli		·	
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	2173	99
		}	(rhodopsin family) protein similar to high-	ļ	1
736	AF132599	Homo sapiens	affinity lysophosphatidic acid receptor homolog) RANTES factor of late activated T lymphocytes-	245	56
.50	1.1.1.2.377	110mo sapions			33
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99.
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus	open reading frame (196 AA)	83	24
741	W03626	musculus	Thursday of CDD XX	110	10
741 742		Homo sapiens Homo sapiens	Human thyrotropin GPR N-terminal sequence. V segment translation product	118 614	40
743	U66059 AF119815	Homo sapiens	G-protein-coupled receptor	2751	100 99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32	448	95
	1		clone HLTCJ63.		
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA	968	65
740	1,04000	 	from plasmid pGCS2232.	(22	1.00
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein	622	100
749	AL022238	Homo sapiens	sequence SEQ ID NO:76. dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	391	87

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus	granuphilin-a	773	41
	1	musculus	, vmaaaa	l	100
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus norvegicus	vasopressin receptor	979	68
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus	netrin 4	2019	89
		musculus		1	
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOVI	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
				t-104	1.62
794	G04072	Homo saniene	Human secreted protein SEO ID NO: X153	1 124	102
794 795	G04072 Y69384	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8153. Amino acid sequence of a 14274 receptor protein.	124 119	100

SEQ	Accession	Species	Description	Smith-	1%
ID	No.) Specific	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Waterman	Identity
NO:				Score	
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CR1 protein.	11963	97
802	P92219	(human)	CK1 protein.	11903) 9/
902	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
803		Homo sapiens			
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
			LAT2	<u> </u>	
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
			clone HOVBA03.	1	}
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
		1	encoded by GenBank Accession Number	J	Į
			L25899		1
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
0.0	}	1201110 0=p10110	gn114 1.	1	1
817	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
818	AF151800	Homo sapiens	CGI-41 protein	1106	95
819	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
820	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
	Y50125			4897	99
823	130123	Homo sapiens	Human glycophosphatidylinositol-anchored protein GPI-122.	4897	99
824	AF156778	170	ASB-3 protein	2675	98
825	AF096322	Homo sapiens		1105	100
823	AF090322	Homo sapiens	neuronal voltage-gated calcium channel gamma- 2 subunit	1103	100
026	1/07070	17.		1540	100
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded	1540	100
007	A D020012	1	from gene 28.	0426	-
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
	1	1	gene 24 SEQ ID NO:147.	1	
00.0	345.4000	 		1.000	ļ. <u>. </u>
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
831 832	X14830 Y71262	Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1.	2540 1002	100
831 832 833	X14830 Y71262 G03873	Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954.	2540 1002 638	100 100 96
831 832 833 834	X14830 Y71262 G03873 AC003030	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1	2540 1002 638 1389	100 100 96 93
831 832 833 834 835	X14830 Y71262 G03873	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein.	2540 1002 638 1389 964	100 100 96 93 87
831 832 833 834	X14830 Y71262 G03873 AC003030	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1	2540 1002 638 1389	100 100 96 93
831 832 833 834 835 836	X14830 Y71262 G03873 AC003030 Y38422 U41557	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich	2540 1002 638 1389 964	100 100 96 93 87 36
831 832 833 834 835	X14830 Y71262 G03873 AC003030 Y38422	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein.	2540 1002 638 1389 964	100 100 96 93 87
831 832 833 834 835 836	X14830 Y71262 G03873 AC003030 Y38422 U41557	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich	2540 1002 638 1389 964 85	100 100 96 93 87 36
831 832 833 834 835 836	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in	2540 1002 638 1389 964 85	100 100 96 93 87 36
831 832 833 834 835 836	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	2540 1002 638 1389 964 85	100 100 96 93 87 36
831 832 833 834 835 836 837	X14830 Y71262 G03873 AC003030 Y38422 U41557	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3.	2540 1002 638 1389 964 85 998	100 100 96 93 87 36 75
831 832 833 834 835 836 837 838 839 840	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943.	2540 1002 638 1389 964 85 998	100 100 96 93 87 36 75 60 67 92
831 832 833 834 835 836 837 838 839 840 841	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ 1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731.	2540 1002 638 1389 964 85 998 1580 1105 255 644	100 100 96 93 87 36 75 60 67 92 97
831 832 833 834 835 836 837 838 839 840 841	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1	2540 1002 638 1389 964 85 998 1580 1105 255 644 2629	100 100 96 93 87 36 75 60 67 92 97
831 832 833 834 835 836 837 838 839 840 841	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1 Human secreted protein clone yc27_1 protein	2540 1002 638 1389 964 85 998 1580 1105 255 644	100 100 96 93 87 36 75 60 67 92 97
831 832 833 834 835 836 837 838 839 840 841 842 843	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717 Y73446	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1 Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	2540 1002 638 1389 964 85 998 1580 1105 255 644 2629 1089	100 100 96 93 87 36 75 60 67 92 97 99
831 832 833 834 835 836 837 838 839 840 841 842 843	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717 Y73446	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1 Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114. Human secreted protein, SEQ ID NO: 6953.	2540 1002 638 1389 964 85 998 1580 1105 255 644 2629 1089	100 100 96 93 87 36 75 60 67 92 97 99 100
831 832 833 834 835 836 837 838 839 840 841 842 843	X14830 Y71262 G03873 AC003030 Y38422 U41557 AL121889 AJ011415 W80398 G00862 G02650 AF036717 Y73446	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Caenorhabditi s elegans Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	acetylcholine receptor beta-subunit preprotein Human chondromodulin-like protein, Zchm1. Human secreted protein, SEQ ID NO: 7954. R29828_1 Human secreted protein. glycine-rich dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) plexin-B1/SEP receptor A secreted protein encoded by clone cw1543_3. Human secreted protein, SEQ ID NO: 4943. Human secreted protein, SEQ ID NO: 6731. FGFR signalling adaptor SNT-1 Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	2540 1002 638 1389 964 85 998 1580 1105 255 644 2629 1089	100 100 96 93 87 36 75 60 67 92 97 99

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	 	ļ	4- A F028040 (DVD2822202)	Score	ļ
848	X99886	Homo sapiens	to AF038969 (PID:g2827207) monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to P34984 (PID:g464305)	963	98
850	AB038237	Homo sapiens	G protein-coupled receptor C51.2	1767	100
851	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by gene 25 SEQ ID NO:305.	1048	98
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 ·	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875 876	Y27572 M15530	Homo sapiens Homo sapiens	Human secreted protein encoded by genc No. 6. B-cell growth factor	573 171	96 56
877	W63681	'Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus norvegicus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
385	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887 888	X92744 Y22496	Homo sapiens Homo sapiens	hBD-1 Human secreted protein sequence clone	375 994	100 94
889	Y41293	Homo sapiens	cn621_8. Human soluble protein ZTMPO-1.	4595	99
390	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
391	AF208856	Homo sapiens	BM-014	1012	99
392	U29195	Homo sapiens	neuronal pentraxin II	2002	98
393	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo sapiens	G0S19-2 peptide precursor	481	100
897	Z68747	Homo sapiens	imogen 38	2018	99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
899	AF225420	Homo sapiens	AD025	734	100

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.			Waterman Score	Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162_1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP-62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93 63
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	1
936	X56385	Canis	rab8	1064	100
937	B08906	familiaris Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 496.	402	70
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRJ	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein sequence SEQ ID NO:110.	587	100
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CG1-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc# AF030433	1466	100
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRPI	632	100
969	AB021227	Homo sapiens	mcmbrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide gated cation channel hHCN4	6295	100
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP80	3348	99
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412- encoded protein.	1553	99
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	172	70
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca sexta	juvenile hormone esterase binding protein	226	32
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
990	AF204159	Homo sapiens	potassium large conductance calcium-activated channel beta 3a subunit	1486	100
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi s elegans	VM106R.1	327	40
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995	AF133845	Homo sapiens	corin	5811	99
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	4999	100
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID NO:212.	725	100
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56.	676	47
1001	AF190167	Homo sapiens	membrane associated protein SLP-2	1747	100

SEQ ID NO:	No.	Species	Description	Smith- Waterman Score	% Identity
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005 ·	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMPI	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	85
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082 1083	L13802 W75098	Homo sapiens	ribosmal protein small subunit	499	80
		Homo sapiens	Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084 1085	G03564 G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1086	AF090942	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8144. PRO0657	88 124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123 5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	59
			clone HTDAD22.		<u> </u>

SEQ ID	Accession No.	Species	Description	Smith- Waterman	%
NO:	No.	1		Score	Identity
1101	AF119851	Homo sapiens	PRO1722	183	72
1101	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1102	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8167. Human secreted protein, SEQ ID NO: 8144.	91	52
1103	X74856	Mus		128	69
		musculus	ribosomal protein L28		
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7432.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
1121	G03714	Homo sapiens	ID NO. 155. Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830			703	94
1123	AF212862	Homo sapiens Homo sapiens	two pore domain K+ channel; TASK-2	442	88
1125	W64469	Homo sapiens	membrane interacting protein of RGS16	191	53
1126	G01361	Homo sapiens	Human secreted protein from clone CW795_2. Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens		88	73
1130	Y32923	Homo sapiens	Human secreted protein, SEQ ID NO: 6186. Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by	525	96
1133	Y91449	Homo sapiens	gene 43 SEQ ID NO:317. Human secreted protein sequence encoded by	542	100
	L	L	gene 49 SEQ ID NO:170.		
1134	AB017908	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	АВ030235	Canis familiaris	D4 dopamine receptor	89	48
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human Secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
		s elegans	cerevisiae zinc resistance protein		
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181,	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
1154	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens	Turnour suppressor protein, p53.	341	87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	99	41
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Hurnan secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus norvegicus	tRNA selenocystcine associated protein	249	62
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	107	71
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens Cl542_2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID NO:175.	1089	89
1200	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	154	57

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.		•	Waterman	Identity
NO:	1			Score	
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	polypeptide. High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB). Extended human secreted protein sequence, SEQ	135	31
1223	Y00278	Homo sapiens	ID NO. 160. Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	УГРІВ	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

SEQ ID	Accession No.	Species	Description	Smith-	%
NO:	No.			Waterman Score	Identity
NO.			ID NO. 160.	Score	
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
	1	norvegicus	Branch receptor microstang protein 2	1	1
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	<u> </u>		gene 27 SEQ ID NO:131.		1
1251	L15313	Caenorhabditi	putative	858	59
		s elegans		<u> </u>	
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate	278	75
1253	W01730	Homo sapiens	reading frame protein. Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-l	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
	10,5.0	Troine daysons	from gene 26.	"	1
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	986	100
	i	, ,	ligand (clone 3TW).		1
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	288	71
	<u> </u>		sequence.	Ĺ	<u> </u>
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
1266	Y13397	ļ	ID NO:2.	<u> </u>	ļ
1267	AF030558	Homo sapiens Rattus	Amino acid sequence of protein PRO334. phosphatidylinositol 5-phosphate 4-kinase	191 859	100
1207	Ar030336	norvegicus	gamma	839	93
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2	552	76
	12.5000	musculus		1 332	/ / /
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
			GTPase-activating protein p135 SynGAP)	}	'-
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthethase (acetate-coA ligase))	1280	100
1274	AF064748	Mus	\$3-12	3523	61
		musculus		<u>.</u>	ļ
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted	643	90
1277	AE146760	IVama anniana	protein.	202	100
1277	AF146760 Y05069	Homo sapiens	septin 2-like cell division control protein	707	100
1279	X59668	Homo sapiens Oryctolagus	Human PIGR-2 protein sequence. aorta CNG channel (rACNG)	281 267	85
1217	7.55008	cuniculus	aura CNO Chathier (INCNO)	207	83
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus	odd-skipped related 1 protein	357	98
		musculus	· · · · · · · · · · · · · · · · · · ·	-	"
1284	U87318	Xenopus laevis	NaDC-2	535	60
1285	AF061346	Mus	Edp1 protein	452	68
1007	1 10000	musculus			
1286	AB030182	Mus	contains transmembrane (TM) region	582	68
1202	A 12505	musculus	i DNS	106	<u> </u>
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
1288	AF254411	Homo sapiens	certain rich are mPNA salising factor CD A1	927	100
	AF084205	Rattus	ser/arg-rich pre-mRNA splicing factor SR-A1 serine/threonine protein kinase TAO1	837 319	100 98
1289	I A HDXA 7015				

SEQ	Accession	Species	Description	Smith-	1%
ID a	No.			Waterman	Identity
NO:	1	1		Score	
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine	468	100
1271	FA 034037	Homo sapiens	deaminase	1 400	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis	ATP-dependent RNA helicase-like protein	636	45
1293	AB010092	· ·	A 1 P-dependent KNA nencase-like protein	030	43
1004	1 5000000	thaliana	<u> </u>	1400	1.00
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22	504	98
			clone HFEAF41.	<u> </u>	
1296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating	459	81
		1	type receptor protein JEG18.	ļ	
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
1305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying	332	98
		1201110 0111111111111111111111111111111	protein SEQ ID NO:1.		1
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus	arsenite inducible RNA associated protein	688	70
	}	musculus	•		
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus	Elf-1	806	92
1316	019017	musculus	C11-1	800	92
1319	U82598	Escherichia	ferric enterobactin transport protein	768	100
		coli		1	
1320	D90892	Escherichia	SORBITOL-6-PHOSPHATE 2-	709	100
	1	coli	DEHYDROGENASE (EC 1.1.1.140)	ļ	ļ
			L /CLLICITOL. & DUOQDUATE		I .
		1	(GLUCITOL-6- PHOSPHATE		
			DEHYDROGENASE) (KETOSEPHOSPHATE		
			DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).		j
1321	W67847	Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41	601	92
			DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74.		
1322	AJ276101	Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein	466	93
1322 · 1323	AJ276101 AJ276101	Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein		
1322	AJ276101	Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21.	466	93
1322 · 1323	AJ276101 AJ276101	Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein	466 504	93 97
1322 1323 1324	AJ276101 AJ276101 Y58628	Homo sapiens Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21.	466 504 1584	93 97 100
1322 1323 1324	AJ276101 AJ276101 Y58628	Homo sapiens Homo sapiens Homo sapiens Rattus	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21.	466 504 1584	93 97 100
1322 1323 1324 1325	AJ276101 AJ276101 Y58628 U91561	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by	466 504 1584 1277	93 97 100 89
1322 1323 1324 1325 1326 1327	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBC174. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826.	466 504 1584 1277 1606	93 97 100 89 100 90
1322 1323 1324 1325 1326 1327	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206 AF151048	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBC174. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826. HSPC214	466 504 1584 1277 1606 1531	93 97 100 89
1322 1323 1324 1325 1326 1327	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206 AF151048 Y10530	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBC174. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826.	466 504 1584 1277 1606	93 97 100 89 100 90
1322 1323 1324 1325 1326 1327	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206 AF151048	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBC174. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826. HSPC214	466 504 1584 1277 1606 1531	93 97 100 89 100 90
1322 1323 1324 1325 1326 1327 1328 1328	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206 AF151048 Y10530	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826. HSPC214 olfactory receptor guanine nucleotide exchange factor sodium dependent phosphate transporter isoform	466 504 1584 1277 1606 1531 657	93 97 100 89 100 90 85 100
1322 1323 1324 1325 1326 1327 1328 1329 1330	AJ276101 AJ276101 Y58628 U91561 AF125533 Y32206 AF151048 Y10530 AF180681	Homo sapiens Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens Homo sapiens	DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE). Human secreted protein encoded by gene 41 clone HPBCJ74. GPRC5B protein GPRC5B protein Protein regulating gene expression PRGE-21. pyridoxine 5'-phosphate oxidase NADH-cytochrome b5 reductase isoform Human receptor molecule (REC) encoded by Incyte clone 2825826. HSPC214 olfactory receptor guanine nucleotide exchange factor	466 504 1584 1277 1606 1531 657 1645 4314	93 97 100 89 100 90 85 100 99

SEQ	Accession	Species	Description	Smith-	%
ID `	No.		-	Waterman	Identity
NO:		·		Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus	protein phosphatase 2C	1931	95
	ļ	norvegicus		1	1
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339 ·	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus	GTP-binding protein	1167	97
	}	norvegicus		1 .	1
1344	AC020579	Arabidopsis	putative phosphoribosylformylglycinamidine	3283	51
	j	thaliana	synthase; 25509-29950	1	
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58	1171	100
			clone HHFHN61.	l	.l
1347	M55542	Homo sapiens	guanylate binding protein isoform l	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel	562	99
	<u> </u>		modulatory subunit	<u> </u>	

TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
1		A		337	1	HWPQAPHRA***GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMRQMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R SCNYTLALILFL
4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

Decidide Seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence		1	hod	ID NO:	beginning		D=Aspartic Acid, E=Glutamic Acid,
1360 137 1361 1361 1362 A 177 12 1362 A 177 12 1363 A 249 355 105 Market 1364 A 1364 A 254 572 201 1366 A 257 425 68 1366 A 257 425 68 1366 A 257 1366 A 257 1366 A 257 425 68 1366 A 258 68 208 Narriery language 1366 A 258 68 208 Narriery language 1366 A 258 68 208 Narriery language 1366 A 258 68 208 Narriery language 15 150 15			1				
1	1		l	i .			
amino acid residue of peptide residue of peptide residue of peptide sequence peptide		nence	1				
Pepide Sequence	Canc		ļ	1 714			
peptide					1		
	1		ļ			Soquence	/=possible nucleotide deletion. \=possible
WSHEGELLQAFRGHIQGGGIRALAHERQAWY TIGGDDSGIRL WILLYGGYRGLGGJGSLLQ VP**ARYTQGCDSGWLLATAGDPYRGPVSL PRGPVSL PRGPVGAAARG**PTFPVLIPAGGSSWSRGL RIVCYGQWGRSQGCPHQHSNCCCGPDPVS WGGAQLELGPAWL			1	1		1	
			 	 -	- soquenee		
	İ			1		1	
RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS WGGQLELGPAWL			ļ	ĺ		Í	
WEGAQLELGPAWL						1	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
WEGAQLELGPAWL					İ		RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
LLTEELH.RGVSIYV.REAQTIPLIVCAL LLVCRRI-\$505CMRAALNDRGI-YQVQLLDGI-V QCLGFVDSDSRKMASTILT 1359			1			!	WEGAQLELGPAWL
LUCREL'SDSCMRAAINDRGLYQVLELDGLV QCLGPVDSDSKRMYSTLT 9	8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
9 1359 A 115 49 186	-		1				LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
1359 A 115 49 186	j	j		j	}	ļ	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
1360	9	1359	Α	115	49	186	
			<u> </u>				
CTLGQIVSQRALSKSLLKHGNTAGKSSITVIA	10	1360	Α	123	2	1249	
		[1	Į			
LEIFRMNDDATQQLVHRTEVVMNNI.SPAWK		[1	[ĺ	
SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK	1	1	1	ļ	ļ		
HDFIGETTSTEKENRGAMEGKQVQWECNIPK YKAKKKNYKNSGTVILNICKIHKMHISFLDYI MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDAINFNEDNPECAGIGGVV EAYQSCFVRAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIHFILVDPGQVGVYSQDMGP DNPGGHFV 11	1	l	l	i .	}	}	
YKAKKMYKNSGTVILNLCKIHKMHSPLDY MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNECAGIQGVV BAYQSCPPKAPTFFITINCHPISSKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRGVILSSSSSSSSLTKDALPSSL KSDSTITISGLVFPFRSLCVNFAKSSVSSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATSVG SGSESPTPFFLTGLQSPPSTRPREGLTTARNS TITLTRDC LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKHFISLLDGNVGPDMSAGFCPNHKAAM VLFDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDTI/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLITMMVY 13 1363 A 249 535 105 WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLMHLFLLGGNA*PHISVEC/CFS TFEIAIKTISFVLYFRYAPPEVLLRSSVYSSPI DWWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL VYLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPV*KPLLVWGVTI *SFFFIESWGTNGGLLSAFTYACVAALCHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN TMVMNRPLCTATVNATNKMGFLNSQVN TMVMNRPLCTATVNATNKMGFLNSQVN TMVMNRPLCTATVNATNKMGFLNSQVN TMVMNRPLCTATVNATNKMGFLNSQVN TMVMNRPLCTATVNATNKMGFLNSQVN TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKK KQTPNSETAPSVCRNLVPENCG THAKFLNKKPNIPKLVIPKLVIPVKAIPTKM AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIBLDQDSVVKLENWTDASRV SKDVFSKPVNIFWALEESVLGWKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIBLDQDSVVKLENWTDASRV VSSLAMKEMLTKTTM	ľ						
MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNECAGIQGVV EAYQSCRPKAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV 11	ļ	ļ	ļ		Ì	j	1
YQPNEYLXAL WAYGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNFECAGIGGVV EAYQSCLPKAFPTTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV 11 1361 A 147 614 9 ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRGTLSSSSSSSSITKDALPSSL KSDSTTITSGLYRAKSSVESSVSSI KILLSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESPTPFFLIGLQSPPSTRPREPGLITTARNS TTLTRDC 12 1362 A 177 12 416 LIPSEPALDSLVDPRVRSRKQFFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNY 13 1363 A 249 535 105 WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLNHLFLDGNA*PHSVECYCPS TTELAKITS*PULYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRFLPFGTSEVDEIFKIC QVLGTPKKVSTLYPKLL 14 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEPVS*KPILVWGYIT *SFFFIFSWGTINGCLLSAITYACYAAICHPILS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLULVIKLYJTVKAJPTKM ALEFILECDONITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKK KQTPNSETAPSVCRNLVFDKCG AGNIFEMTCKVSSKNIKSPRYSVLIMAERPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV GDLSSPNETKYIISLDQDSVVKLENWTDASRV VSSLAMKEMLTKTTM 1367 A 298 68 208 RKRTNNPIKLDKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVQQAGL	1	}	ŀ	ļ			
GARIPPEYTOSHOFANIFIEDNPECAGIQGVV				}			, , ,
BEAYOSCEPKAPITTGPTNICPHSSRXVAKFRR SEGN*HQGRAFAIIFILVDPQQVGVYQDMGP DNPGGHFV		1					
SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV	1			1			
DNPGGHPV							
11			i	Ì			, , ,
NTTSSRPASSRGTLSSSSSSSSLTKDALPSSL KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSI KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSI KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSI KSLLSSSVKYLE*KRTSCCFPDSSESKLSQLS DERVSMGTSSRKPTNSSSSLGALKMSATSI*G SGSESPTPFFLTGLQSPPSTRPEPGLTTARNS TTLTRD	11	1361	A	147	614	9	
	1		1			-	
DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESTPFFTLTGLQSPPSTRPREPGLTTARNS TTLTRDC 1362		ļ	İ	[
SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS		1	ļ				KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS
1362 A 177 12 416 LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY 13 1363 A 249 535 105 WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL 14 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*PFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACVAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL		1	•	}			DERVSMGTSSRKPTNSSSSLGALKMSATS*G
1362 A 177 12 416	1		ŀ				SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS
DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM	L			<u> </u>			
VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY	12	1362	A	177	12	416	
ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNYY	1	1	Į.				
NILTMNYY	1	ſ	[[
1363		ļ .	l				1
MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFFGTSEVDEIFKIC QVLGTPKKVSTLVPKLL 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCK VSSKNIKSPRYSVLIMAEKPV GDLSSPNETK YIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	<u></u>	1000	<u> </u>			100	
TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL 14 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITVKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	13	1363	A	249	535	105	
DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL 14 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECQQNITVKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCK VSSKNIKSPRYSVLIMAEKPV GDLSSPNETK YIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	1	ł	}				
QVLGTPKKVSTLVPKLL	1	1	ł	į l			
14 1364 A 254 572 201 YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITVKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCK VSSKNIKSPRYSVLIMAEKPV GDLSSPNETK YIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	İ]]	ļ			
FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN 15 1365 A 257 425 68 THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	14	1364	 	254	572	201	<u></u>
*SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN	14	1304	^	2.24	312	201	
TMVMNRPLCTATVNATNKMGFLNSQVN 15	1]		}			
1365	İ		1				
AIEFLLECDQNITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIK WQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	15	1365	<u> </u>	257	425	68	
TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG 16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	"	1,303	1 ^	231	723	1 00	
KQTPNSETAPSVCRNLVFDKCG		[
16 1366 A 263 104 481 FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL		1	1				
SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	16	1366	A	263	104	481	
AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL		1300	,	203		.01	
GDLSSPNETKYIISLDQDSVVKLENWTDASRV 17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL		1	l				
17 1367 A 298 68 208 RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL							
VSSLAMKEMLTKTTM 18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL	17	1367	A	298	68	208	
18 1368 A 300 904 1 LVVGITGTRHHARVIFIFLVETGFPHVGQAGL		1	-				1
	18	1368	A	300	904	1	
		1	1	[ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHLLSHIWKDVQNILLK
20	1370	A	304		1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW
21	1371	A	326	799	1587	ALVRSRGG GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	Α	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ
25	1375	A	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	Α	397	383	165	EVKNTNPFIFSGTNI.TIWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIVVFLYANDEQLEIEMNK IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL
28	13/8	A	408	14	427	ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/I/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRPANF*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	Α	474	125	471	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825		KSACSFICSEEQPASPSPLKPGTYASETRPRDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR APGGSAGS*\GLPSAGGSRGRKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRQE*P PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS* GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCOFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	. 512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRPVVPFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seg-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchec	į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice		}	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Į			peptide	Sequence	/=possible nucleotide deletion, \=possible
1	l		1		Ì	nucleotide insertion
ļ	 	ļ	 	sequence		l
1				į		GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
	l		1	ľ		GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
1	1					VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
	J		ł	j	}	HGA\GLPGVRVHTQRVHIH*GAG/GCQTPRPR
1	l		ļ			LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
ļ.				1		ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
l	i	i	i	İ	i	NHQAVGLEASGALQAGHRDELPTMVQLLDH
	<u> </u>		<u> </u>			SPDYPLKGRPHAP
37	1387	Α	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
	l		Į.		Į	WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
						RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
1						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
			1			FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
		İ				AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
ł	1		l		ł	SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
		١.			ŀ	\DAEFVFSVLVRESKASAVGDDDKVYYFFTE
1	ļ	1	ļ			RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
1	1	İ				TSEEGKEG*EVPV*LPVSPPLPRPLOKMLDYL
	i	1	ľ		•	KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
l	Ì	l	ľ			QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
	l					INIIICTVDYLLRLQESI
39	1389	A	767	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
1				_		RNLSVDGKNVDMAGFIANNGTREGCAARRN
İ			1			FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
		[KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
İ	ŀ	1				VRGLHIOVROPLVVYAAFTVDSHRPLOETVL
ļ.		1	ł			RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
		1				ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
j	ŀ		j			PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
J	}]]			LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
	l		1		ļ	TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
1	1					SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
	!	ĺ	ĺ		ļ	QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
1 40	1390	1 ^	801	09	399	POITYKFNA*O\VILKFTW*E*GAKITILRKNKL
	1					RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
	ļ		İ		}	1
1	1201	<u> </u>	926	7	106	VNKSVVLVQVTIM
41	1391	A	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
L	1.205	 	1000	<u> </u>	 	SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	Α	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIIHWIES
{	İ		1			KASFGDECSHHAYLHDQFWSYWNSLKHRTW
j	l	}				QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
1	1]	1			FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
				<u></u>		KACFPTNIVTL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
	1					LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
L		}				QRPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
٠.	1		[SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
1	ĺ	İ		Ī		PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
	1					QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
 		1				PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
-		1 -				QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
	· ·] .			WLSMSMGK
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
"	'550	l ^^	'''	•	. 500	EIOKYMRT/DO*CVTHDISLYIVTKLALIFLIPR
			(,		VFI.FHQLNIT**CI.HFFTMTTFIAIPFSFLFLGR
	L	L	L		L	TEST TIQUETTS CONTENTIAL TELATIFICATION

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRRENRLNPGGGG
						CSEIAP\CTPAWVTQRDFFRKKK
48	1398	A	963 967	216 466	308	HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC PRKRESWWGERLP/PRGFPPAAEDAPAPGWK
49	1399	A	90 <i>1</i> 		1	GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPITAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGQYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSIRPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA/LYTWHVLEQLEILRQINQQSHGPG
56	1406	٨	1044	5	429	SVLTLQTRSPSKPLSTRKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchee	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciicc]	į	"	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Josquesse	/=possible nucleotide deletion, \=possible
	į	ĺ	1	sequence		nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
٠,	1407	^`	1050	l ''	130	MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
	Ì	Ì		ĺ		TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
	Į.			ì		ADDTHPARLOGPTLRSQPMGPLKHKAFEERA
	ŀ	Ì				NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
50	****	^`	1050	250	117	PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
39	1707	\ ^	1004	'	123	KGSSSLNNHORIHTGEKPYKCNECGRAFSQC
				ļ		SSLIOHHRIHTGEKPYECTQCGKAFTSISRLSR
	1	1	ļ		1	HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
			l	İ		TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
OU	1410	J A	1000	204	419	LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
		ł			l	EGQEGLPLTV
(1)	123.	 	1070		202	RHSRAHLCOPFHLVMRDLLOLGODIPOGCHY
61	1411	Α	1079	3	383	LEENHLIHRDIAARNCLLSCAAPTRAATIGDF
	ļ			1		
	j	1]	l	j	GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
		l				l ·
70	1412	ļ	1000	-	0.50	TN VVEFLWSRRPSGSSDPRPRRPASKCQMMEER
62	1412	A	1080	1	859	-
	1			1	ł	ANLMHMMKLSIKVLLQSALSLGRSLDADHA
]]			J	PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF
						GPLELVEKLCPEASDIATSVRNLPELKTAVGR
	1					GRAWLYLALMQKKLADYLKVLIDNKHLLSE
	1					FYEPEALMMEEEGMVIVGLLVGLNVLDANL\
	ĺ	{	1	ĺ		CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
	,	l	1		}	HERITDVLDQKNYVEELNRHLSCTVGDLQTK IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
				1		REONELIR
-	1413	 	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
63	1413	Ā	1083	4	013	HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI
	[1	ĺ	[[HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
						EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
		ļ			1	YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
	i	l		i	į.	DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
					ļ	KAFNHTSICCRHKKN
64	1414	 	1084	946	 	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
04	1414	A	1064] 340	1 1	i
]		SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
	Ī	1		1	ſ	LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
		1	}	j]	TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
	!				1	FFSWLTTGLTTQQRTAIE\NATVAFF\LQCI\SC
	1	1		1		HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
			1	1		S\GFIR\RLFLQLMLEDEK\TMFLQSPCPLYKG
		l	1	[ţ	RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL
	İ	1		1	i	SDVLDRVSDTPSITAKLISKOKDDKKKK
66	1415	-	1007	103	224	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
65	1415	Α	1087	103	324	
				l .		LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA SVALHKLSNALV
	1417	 	1005	 	402	SVALHKLSNALV HETCSVTHIVSFSLPFLNPSHPASTPGHTENEO
66	1416	Α	1095	3	493	
	1	1		1		PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
	1	l	1	l		TLPVAAAFTETVNAYFKGADPSKCIVKITGE
				1		MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
	1	ŀ		I		HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL
/2	14:5	 	1000	L	1200	MTHLK
67	1417	Α	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
	l	ļ		ì	ł	PYYFLLDLCCSDILRSAICFPFVFNSVKNGST
	1	1	I	l		WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	[914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ļ	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
})]]	peptide)	/=possible nucleotide deletion, \=possible
				sequence	ļ	nucleotide insertion RYL
68	1418	A	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
60	1410	^	1100	1 .	1320	YEREGMODWKTASGOSEEATOOSSOKPOPH
	1			ł	i	YTTYOSSSFLKYSSESHLLAWRENSSEGSFOF
	!					PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
		i			ĺ	HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
}	1		•	ļ ·	ļ	RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
İ						SPAALAPRAARGGSRAAALAGAEAEEPLRTL
				ļ		APRPTRAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
						RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
					}	APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
	}					GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
						SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
	ľ				}	HQNIHPENQELVRVLREQLQTEQDAPAATRQ
1						QFYTDMYCPICLHQASFPVETNCGHLFCGSLT
						PNSIW
69	1419	Α	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
	l					FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
İ						LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
ļ						TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
						RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	Α	1111	698	23	ALRRLHYVRATKV\FLSFRRPFWREEHIEGGH
						SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
				1		AFAGLSREEALRLALDDVAALHGPVVRQLW
1					!	DGTGVVKRWAEDQHSQGGFVVQPPALWQT
.				ļ .	,	EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
					ļ ,	KSALRAAIKINSRKGPASDTASPEGHASDMEG
						QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL QNTTHTRTSH
71	1421	Α	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
' · '	1.47.1	4.5	****	-	303	PPGPPEQAGLSQFHLEPETQNPETTEEIQSS/LQ
					l	QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
						EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	Ā	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
				-		GOKVAIKIVNREKLSESVLMKVEREIAIL\RLI
						EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
						QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
		1				GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
		l				HSYSICHROLKPENLLLDEKNNIRIADFGMAS
		ł				LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
						RADMWSCGVILFALLVGALPFDDDNLRQLLE
		,	j			KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
		<u> </u>				LSLEQIQKHPWYLGGNFIS
73	1423	Α	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
\						FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
						MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
						SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
						LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
			ŀ			GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
				İ		TAGLNVAAEGARARDMPAQAWDLVERMKN
						SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
-	1404		1120		400	HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
						DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
		1				VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
)				:		EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
75	1425		1147		413	AADPAPVHTTAHPKGA
'3	1423	A	114/	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR
			I			1110DECORY 11 400HUM ELDEWA I DOGUN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ DDESGQKKLHGLQAILVHEASGTTAITATAT
76	1426	A	1155	38	410	GYQESHLSSAR PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVQYIMNCMLWK KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
77	1427	A	1162	526	350	T RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
78	1428	A	1171	I	1293	LRGNLLRVFRQVEKVQEENKWQSPLED MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYQPCDDMDCLSDRCKILQVFDDFIFIFA MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PP\YYQPEEDDEMPFICSLSGDNGIMGCHEIPP LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	i	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI CHKVIMGAGKLPRHMRTHTGEKPYMCTICE VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82	1432	Α	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC WGRGHGCGQEALSTSHGYHLFCALLTGFLFA SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES GACVNQVTVDSITPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT GRSPCPSLPGTTRTNSLL
86	1436	Α	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC NLGFFLEGHAVLTCHAGSENSATWDFPLPSC RADDACGGTLRG/AEWHHI.QPPLPLG/ATKN

NO. of much entitle estate sequence No. of much estate sequence No. of much estate sequence No. of much estate sequence No. of much estate sequence No. of much estate sequence No. of much estate sequence No. of peptide sequence No. of	SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
nuclectide seq unnec co						1	
uence USSN grote of peptide corresponding leader in the state of peptide residue of peptide sequence solve of peptide sequence sequence solve of peptide sequence se	1	I .				1	
Sequence 9,496 1437 216 226 964 1438 A 1218 1 234 1235 1438 A 1227 2 349 1440 A 1227 2 349 344 A 1245 3 1937 1441 A 1245 3 1937 1441 A 1245 3 1937 1441 A 1245 3 1937 1441 A 1245 3 1246 5 562 444 5 562 444 5 562 444 5 562 444 5 562 444 5 562 444 5 562	eotide		İ	USSN			, ,
1437 A 1216 226 964 Maniposid residue of peptide p	1		1	1	E .		
amino acid residue of peptide residue of peptide residue of peptide sequence per per per per per per per per per pe			ł	L		•	
Persidue of peptide sequence			l i				
Peptide sequence Independent Independ	ţ	1	j		residue of		
	1 .				peptide	ļ · •	
NADCTWTILAELGDTTALVFIIPQLEDGYDFL	(ĺ	İ			, .
87					50420000		
1437	1		İ		l		,
	87	1437	A	1216	226	964	
] "	1	١	1.2.0]] "	, : = :
						Ì	
			1				
			ļ				
BENIKKLADOFLEGOKQĒTOKIOSNOŪKELDI NNOVPKNIPK VAENVADKNEPSSNHIPHG	i		ł		Ì	İ	
NNQVYKNIPKVAENVAJKNEEPSSNIPHPIG 88	1				Į.		
88			!		,	ļ	
RVKSAGAWIIHPYSDFRFYWDLMLLIMVĞN LIVLPVGITFKEENSPPWITVLSDTFFLLD LVLNFRTGIVVEGGAELLAPRAIRTSYLRTW FLVDLISSIPVDYIFLVELEPRLDAEVYKTAR ALRIVRFIKLSLLRL 89 1439 A 1223 1 743	88	1438	Α	1218	<u> </u>	534	
LIVLPVGITFEKEENSPPWIVFINULSDITFELLD LVVINFETGIVVEEGAEILLAPRAIRTRYLRTW FLVDLISSIPVDYIFLVVELEPRLDAEVYKYAR ALRIVRFTKLISLLEN. 89	00	1430	^	1210		334	
LVINFRTGIVVEEGAEILLAPRAIRTRYLRTW FLVDLISSIPVOVIFILVVELEPRLDAEVYKTAR ALRIVRFTKILSLIRL 89			['			ļ	==:
89 1439 A 1223 1 743 MGFDEVFMINLRRRQDRRERMLRALQAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNLINPDTI-SLLIAEMKTVV APMIDSRAAYSNFWCGMTSOGYYKRTPAYI PIRKRDRRGGFAVPPWISFILDIRKAASRNI VAFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLVFLRAHSITLQDEAESFMHVOJ PIRKRDRRSPSAQSMAVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMITTPFRIVIDAGF GPWDFKFLCRYTSVLFYANMDTSIVVLGITI/ YDRYWKVVRHL/WDSWMTGUSFTRVYLLG IGSSDVRAPQRSEIGAESFSRNVASQAYNLT SALTPILTRSRVLNEEPLTLAGRSRAPANLSD VOQLIFLVDSNPFPFGVILAKGGGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSEIGAESFSRNVASQAYNLT SALTPILTRSRVLNEEPLTLAGRSRAPANLSD VOQLIFLVDSNPFPFGVINTYTSTKVASMAF QTQAGAQIPIERLASERATIVKVPNNSDWAAR GHRSSANSVVQPQAFVQAVVTLDSSNPAAV LHQLNYTLLDGRYLSEEPEPYLAVVLHSEPR PNEHNCSASRRIRESLQGAHRPYTFFISPGT RDPVGSSYRLNLSSHRRALEVSVGLYTSLC QYFSEEDVVWRTEGLLFLEFSFSRQAVCLTR HLTAFGTSLVPPSHIRGALEVSVGLYTSLC QYFSEEDVVWRTEGLLFLEFSFSRQAVCLTR HLTAFGTSLFVPSHIRVSLDGNAFHRNSLDIPQIATP HSI_GSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVADBUGGTAFHRNSLDIPQIATP HSI_GSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVADBUGGTAFHRNSLDIPQIATP HSI_GSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVADBUGGTAFHRNSLDIPQIATP HSI_GSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVADBUGGSTAATCVLLICL FIGANAVWGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYVALALFLRRNSRSV GWGWGGGSTGNGGAWGTCVLLICL FIGANAVWGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYVALALFLRRNSRSV GWGWGGGSTGNGGAWAPCEPPLSSAAAR GKGVHQRILGKGQHT 92 1442 A 1246 5 562 VFDEENLINLENDFIRERIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEFYPMMRRAFETVAIDLDRIGKKN SILLOKFOKDLNITGPWNQENEILKQIVKH PGRKASCSTAGSGSGCLPPASSPMVSSAHNPN KAEIPERKRKDSTSTPNNLPSSMMTRRNTYVCT	1 :		ł i			ĺ	,
ALRIVRFTKILSLIRL	1						
89 1439 A 1223 I 743 MGFDEVFMINLRRRQDRRERMLRALQAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNLLIMPDTLSLLIAENKTVV APMILDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGGFAPPWTSFLIDLRKAASRNL AFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLFVPLRAHSILQDEAESFMHVYS LEVMYPSSPSSAQSMAVSADHIGLVISYL LEVMYPSSPSSAQSMAVSADHIGLVISYL SAMPWKVARHLWDSSPSSAQSMAVSADHIGLVISYL GARLVWFGKLILAKGGHIGGISVIL LGARLVWFGKLILAKGGHIGGISVL LGARLVWFGKLILAKGGHIGGISWL LGASDVARPORSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGPSRAPANLSD VVQLIFLVDSNPFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPAFVGAVVTLDSNPFAAV LHQLNTYTLDGRYLSEPEPYLAVVLHSPPR PNEHNCSASRRIRPSLQGADHRPYTFFISPGT RDFVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWNTFGLLPLETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVPPETTADVNYIVML TCAVCLVTYMVMAALHLDQLDASSGRAP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVJDSSGHRHLDGDRAFHKISDIFQAITT HSILGSMWRIRVWHDNIKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCTRIQRATCCVLLICL FLGANAWYGAVGAVSAYSTGRVSRLNPLSV DTVAVGGLVSSVVVYVTLAILFLFRRNSRSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADNFVTAMLSKLRFEVFQFGDYIRRG AVGKMYFIGHGVAGVIKSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVAIDDLDRIGKKN SLLAKFEVFQFGDYIRRG AVGKMYFIGHGVAGKGLNTIGSVFINSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLVSLSVD NFNEVLEEYPMMRRAFETVAIDDLDRIGKKN SLLLQKFQKDLNTIGVFNNQENEILKQIVKH TVPPPPGGPSPAPLHPKRSPTSTGAELKEERL PGRKASCSTAGSGRGRIPPSSPMYSSAHNPN KABIPERRKDSTSTPNLPSSMMTRRNTYVCT							"
CRL VEAVDGK VGMLTRSNAAPGRHLÄMLET LVVVAPRFVDADNLILNPDTLSLLIAENKTV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLLIDLRKAASRNIL AFYPPHPDYTWSFDDIVFAFSCKQUAEVOMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSPJPYLKNIVVADLIMTLTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT' YDRYYMKVVRHL/MVSMTGIJSFRRYVLIG LGARLVWFGKLILAKGGHGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRWVASQAYNLT SALTPLITRSRVLNEEPLTLAGRISRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAGJPIERLASRATIVKVPNNSDWAAR GHRSSANSVLVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSPER PNEHNCSASRRIPESLQGADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSGLYTSLC QYFSEEDVWRTEGLJELFESPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLQDLDASRGRAD LYGYDSRSGRIRLDGDRAFHINSI.DIFQLATP HSILGSMWRIRW WHDDKSLGPAWBLGPHIVRD LYGYDSRSGRIRHLDGDRAFHINSI.DIFQLATP HSILGSMWRIRW WHDSKLGPAWBLGPHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASRRYPTPSNAALLRTRRLLVAELQRGF FDKHIWLSJWDRPPSNAALLRTRRLLVAELQRGF FDKHIWLSJWDRPPSNAALLRTRRLLVAELQRGF FDKHIWLSJWDRPPSNAALLRTRRLLVAELQRGF FDKHIWLSJWDRPPSNAALRTRRLLVAELQRGF FDKHIWLSJWDRPSNAALRTRRLLVAELQRGF FDKHIWLSJWDRPSNAALRTRRLLVAELQRGF FDKHIWLSJWDRPSNAALRTRRLTVAELDRGRKN GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLGKGQHT FLANDPMTYTAMLSKLRFEVFOPGDYIREG AVGKKMYFIDHGVRGKWKH TOMPSTALTRRLLVAELQRGF FDKHIWLSJWDRPSNATHTNYTVCT 1442 A 1246 5 562 VFDEENILENDENDEN	80	1/30	_	1222	-	742	
LVVVAPRFVDADNILILNPDTI.SLLIAENKTVV	85	1437	^	1223	1	743	
APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL VAYYPHPDYTWSTPSDDIIVFARSCRQAEVQMY VCNKEEYGFLPVPLRAHSILQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMILTFPFRIVHDAGF GPWDFKFILCRYTSVLFYAMDTSIVVLGLITI YDRYWKVVRHLWDSWMTGINSFTRVYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLVAEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEFPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFYPBHAVNTIVML LGYDVSSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKIRVWHDNKGLSPAWFLQHIHVRD LQYDSSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKIRVWHDNKGLSPAWFLQHIHVRD LQYDSSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKIRVWHDNKGLSPAWFLQHIHVRD LQTARSTFFILVNDWLSVETEANGGLVEKEVL AASKASRRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWGAVGADAYSTGRVSRINPLSV GWGWGPGSTGNGAWASAPCEPPLSSAAAR GKGVPGRLGKGGOTTAHSKERKV GWGWGPGSTGNGAWASAPCEPPLSSAAAR GKGVPGRLGKGGVGYTKSNSERMKLTDGS VFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFGKDLNTIOVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEPPERKNDSTSTPNLPPSSMMTRRNTYVCT							
PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNI. VAFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTI,QDEAESFMHYQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHLWDSWMTGIJSFTRYVLLG LGARILWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFISRAPANLSD VVQLIFLVDSNFPFGYISNYTVSTKVASMAF QTQAGAQPIERLASBRATIVKVPNNSDWAAR GHRSSANSVVQQAFVGAVVTLDSSNPAAV LHIQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQCADHRPYTFFISPGT RDPVGSYRLNLSSFRWSALESVGLYTSLC QYFSEEDVWWRTEGLLPLETSPRQAVCLTR HLTAFGTSLFVPPSHIRFWFPEPTADVNYIVML TCAVCLVTYMYMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LVGVDSRSGHRHDDGMAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVADULSVETEANGGLVEKEVL AASKASFRVPTSVAALLRFRRLUVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFV1AMLSKLRFEVFGRGDYIREG AVGKKMYFIGHAGVGVTHTSKEKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLDKTGVFNNQENEILKQIVKH PGRKASSCSTAGSGRGLPPSSPMVSSAHNPN KAEIPERRXDSTSTPNNLPPSMMTRRNTYVCT		•	ľ				
1440 1227 2 349 NKTSPIPLYWSFDDILYAFSCKQNAEVQMY VCNKEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL 1440 A 1227 2 349 NKTSPIPYLKNIVVADLIMTLTPPRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLWDSWMTGISFTRVYLLG LGARLVWFGKLILAKGGHGGISWL LGARLVWFGKLILAKGGHGGISWL LGARLVWFGKLILAKGGHGGISWL LGARLVWFGKLILAKGGHGGISWL LGARLVWFGKLILAKGGHGGISWL VVQLIFLVDSNPFPFGYISNTTVSTKVASMAF QTQAGQPIRELASERATIVKVPNISDWAAG GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PIEHNCSASRRIPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HILTAFGTSLFVPPSHIRRVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQATT HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVANGLVSKEVYLLGLAGF FDKHIWLSIWDRPPRSCFFRQRATCCVLLICL FLGANAVWGAVGDSAYSTGRYSKLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVORLLGKQOHT S62						1	, -
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVYRHL/WDSWMTGI/SFTRVYLLG GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVYRHL/WDSWMTGI/SFTRVYLLG GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVYRHL/WDSWMTGI/SFTRVYLLG LGARLVWFGKLLAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLITLAGF)SRAPANLSD VVQLIFLVDSNFPFFGYISNYTVSTKVASMAF QTQAGAQPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHIQLNYTLLIGBYSLSEPEPYLAVYLHISEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPETTADVNIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKFYEIL VKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVANDWLSVETEANGGLVEKEVL AASKASFRYPTPSNAALLFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQOHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIRRG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLITKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKRQKDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGGRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSSMMTRRNTYVCT	!						i I
LEWMYPSSPSSAQSMA,VVSADHIGLVISYL	j j						·
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG LGARL/WFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGASFSRNVASQAYNLT SALTPILTRSRVI.NEEPLTLAGFSRAPANLSD VVQLIFL/DSNPFPFGYISNYTVSTK/ASMAF QTQAGAQIPIERLASBRAITVK/PRNSDWAAR GIRBSANSVVPQAFVGAVVTLDSSNPFAW LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSVRLNSSHFRWSALEVSVGLYTSLC/ QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFYFPEPTADVNYIVML TCAVCL/TYMVMAAILHKLDQLDASRGAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLIGDGRAFHRNSLDIFQIATP HSLGSMWKRVWHDNKGLSPAWFLQHIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALRFRRLLVAELQRGF FDKHTWLSIWDRPPRSCFTRIQATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLALIFLFRMSRSKV GWGWGFGSTGNGAWASAPCEPPLSSAAAR GKGVHQRLLGKQHT 92 1442 A 1246 5 562 VPEDENILNELNDPLREEINVFNCRKLVATMP LFANADPNFVT/AMLSKLRFEVFOPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVADRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 17PPPFGGFSPAPLHFKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERKKDSTSTFNNLPPSSMMTRRNTYVCT		'					
GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGIJSFTRVYLLG LGARL/WFGKLILAKGGHGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSIKVASMAF QTQAGAG/IPERLASBRAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRPVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDRSGRHHLDGDRAFHIRNSLDIPQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTSMALLRERRLLVAELORGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQOHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADIPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEPYPMMRRAFETTVADDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPPERRKDSTSTPNNLPPSNMTRRNTYYCT		1440	_	1227		240	
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNIT 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNIT SALTPILTRSRVLNEEPLTLAGPSRAPANLSD VVQLIFLVDSNPFFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLPLEETSPRQAVCLTR HLTAFGTSLPPSHIRPTYPEPTADVNIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWPLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSNAALLRFRRLLVAELQRGF FDKHIWLSJWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEIVNFNCRKLVATMP LFANADNPTYAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEPYPMMRRAFETYADDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 30	1440	A	1221	2	349	
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAYVLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASSGRAIP FCGQRGRFKYEILVXTGWRGGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQQHT 92 1442 A 1246 5 562 VFDEENILNELNDFLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCTAGGSRGRIPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1						
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANISD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPPADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTSINGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSRAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	j j					,	,
SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD VVQLIFLVDSNPFFPGYISNYTVSTIKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GTRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEFEPYLAVYLHSEPR RPNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVVD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFIEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMYSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	01	1441		1245		1022	
VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHIQLNYTLLDGRYLSEFPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGGYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	91	1441	A	1245	3	1937	1
QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVIRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSKLNPLSV DTVAVGLVSSVVVPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPPIVIAMLSKLRFEVFQPGDYIREG AVGKKMYFQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
GHRSSANSVIVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLYTMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPNSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMIRRNTYVCT	i !						-
LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVILAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTIGVFTNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	l i						
PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRINPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENLINELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	,						
RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFK YEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT VFDEENILNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSI'GRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVO NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	ļ l						
HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	 						
LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							•
AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	[,
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 [ĺ	. 7.
FLGANAVWYGAVGDSAYSTĞRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	i			ļ		ì	,
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	<u> </u>	1440		-12:2		560	
AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	92	1442	A	1246	o	202	- 1
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							1
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	}						•
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			i				
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT						ļ ļ	
PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	93	1443	Α	1249	180	901	
ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS	}		- 1				
	لــــا						ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
94	1444	A	1261	3	385	TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP RQEDHLSPGGRGCSEL KFSQWGLTKPKLSNASP/WISLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR
95	1445	A	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILKLARH STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA
96	1446	A		1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRKQKEQSGGEEKDED EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC DGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97		A	1295	2	2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
98	1448	A	1304	118	453	SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
99	1449	A	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP TTCTVLLLAESEGERERWLQVLGELQRLLLD ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ QLTLSPSAGLLVVLCGRGPSVRLFALAELENI EVEVPKIPESRGCQVLAAGSILQARTPVLCVA VKRQVLCYQLGPGPGPGPWQRIRELQAPATVQ
						SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL GAGLVPEELPPSRGGLGEALGAVELSLSEFLL LFTTAGIYVDGAGRKSRGHELLWPAAPMGW GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE EQQKQQRREMLKDPFVRSKLISPPTNFNHLV HVGPANGRPGARDKSP
100	1450	A	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA LNDSMINETARDAARVQVASTLSVLVGLFQV GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\ PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD HGHPPARMSIFSR
102	1452	Α	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN \DSCLKQKARRLTILLL
103	1453	Α	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	į	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}]	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	}]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	sequence	/=possible nucleotide deletion, \=possible
	l	ŀ	l	sequence		nucleotide insertion
						FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
				Ì		FTYNVGALGGALAPIIGALIAQRLDLGTALAS
		 	L			LSFSLTFVVILRNRRPGKSLVR
109	1459	A	1402	15	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
			İ		ļ	WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
İ		Į.		[[RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS GVYCCRIEVPGWFNDVKINVRLNLQRASTT
110	1460	A	1421	3	350	HEDLSSLLTRGSGNQERERQLKKLISLRDWM
i		\	1421	1	330	LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
!					1	FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
İ	l	1		ł	ł	CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD
İ		}				QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS
İ		1				TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
112	1462	A	1434	46	372	SDSLLFSQDSKLATTS TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR
112	1402	A	1434	1 40	3/2	SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC
				<u> </u>		STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
: 		ł	}	ļ		MSSSTTSSTTSTF
113	1463	A	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
	Ċ					MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR
						GIFLTFLLSNFLIVCVLLFYVSFYLFQSCINFVL
114	1464	Α	1463	1	396 .	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ
	j	J	J			QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP
	}	ļ	j			ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS
	•					EDP*KNA*LKQMHAATTHWQQHQQHQVGC OYHGIMQ
115	1465	Α_	1464	291	2	AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG
					_	GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
	ļ]				MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
		<u> </u>				NYCN
116	1466	Α	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ
]	•				YWTKYQVWEWLQHFLDTNQLDANCIPFQEF
						DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC
117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
	1107	'	1475	•	301	VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
	ļ	İ				QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN
						NPEEELASDPNNEESL*RPWALEDFEIGRPLG
		<u> </u>				KGK
118	1468	Α	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
		[PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL
				Ì		QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS
		ľ	[<u>'</u>	H
119	1469	A	1486	1	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDOL
		1	[NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE
		[[KALTKFLKWVNWDLPQEAKQALELLGKWK
						PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
100	1100	<u> </u>	1405		000	DDEDLLMYL
120	1470	A	1497	3	999	MGESPAV*GYFVLAGMNSAGLSFGGGAGKY
		[}			LAEWMVHGYPSENVWELDLKRFGALQSSRT
						FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP
			l			PDKDLLALEQSKTFYKPDWFDIVESEVKCCK
					·	EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS
						NDLDVPVGHIVHTGMLNEGGGYENDCSIARL
						NKRSFFMISPTDQQVHCWAWLKKHMPKDSN
L	<u> </u>		L			LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
						GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT WDPGTDTALGWSKQPSQSYTLFES*VGSGYII DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA AKHGHSPAVQVLLAQWQDINEMNEKQQTPL HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF HLYEERGGVIDITAWDKDAGKRDDFIGRCQV DLSALSREQTHKLELQLEEGEGHLVLLVTLT ASATVSISDLSVNSLEDQKEREEILKRYSPLRI FHNLKDVGFLQVKVIRAEGLMAADVTGKSD PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL *VALVWKKFQTQSLRLSDLHRKSHLWRGIVS ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRPLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	A	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA LLCVWALSLVIYIGPLLGWRHPAPEDETICQI NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV AKTE
130	1480	Ā	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE AIRYCHGCGVAHRDLKCENALLQGFNLKLTD FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TDIPKMLWQQQKGVSFPTHLSISADCQDLLK RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
132	1482	A	1656	150	48	LSNKVGGESKPKKKK LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV
133	1483	A	1660	3	406	VDAQ RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSPRDRDRER*RITRERERERDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	Α	1678	525	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
137	1487	A	1680	•		AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSQKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHPSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST PEEFSRRLKERMHHNIPHRFNVGLNMKATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRGQQGWDRK YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

NO: of No: of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decidide Seq		, ,			ľ		
Sequence							
149			ŀ		1		
148					1		
		ļ	ŀ				
Peptide	ļ	1	1	ļ			
			•	ļ	peptide		
148	!	İ	}	ļ	, , ,	ļ	, · · · · · · · · · · · · · · · · · · ·
NSLEVIEIQARSSAGTPARAYLIDIPNERYLLOGFA		-	i		 	 -	
ISSGAIYLASSYONKR.PICCKGNI.VKESTGT	i	i .		ĺ			
HHRGPSTSRR*PASPLEOTYGOGRAFLOGRAFL 188	Ì		ļ .				
1488			ĺ	1	ĺ	ţ	
	138	1488	Ā	1686	2	526	
PGGPCNSCP CSAPESGGROPLPGPGTGGCV CVTRGCOTTARTAAAAAPCPAGRAPPGGA						1	
1489		ŀ	ļ			1	
1489	l	ł	i			ł	
139		1	Ī				
1489	1	ł	ì				, ,
	139	1489	A	1693	3	376	
140	1	1	}				
140	1			ļ		į	
140	· '	l	ŀ				
HSLRIPFKPEFIGIAVTIDERCIDASLREFSNFIL KNLEEQREIMVS*EGKLISQLSRGKKIWIWK LVLVEVVKHLSLGTVHCNGKMRPEP 141	140	1490	A	1704	3	376	
141		`	1		_		
1491 1491 A 1743 I 362			}	J			
141	1			1			
DKLELELVLKGSYEDTQTSFLGTASAFRFHY MAAL*TELSGRLRSSKSNGWIGDNSTGYLTV PLRPLTIVKEVTMDVPAPPNVRGLNWMG 142	141	1491	A	1743	1	362	\$
MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV		1	-	••• ••	•	302	
142	ļ			1			,
142		ĺ		1			
LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF GLQLQGGIFATETLSSPLVCFIEPDSPAERCG LLQVGDRVLSINGIATEDGTMEEANQLLRDA ALAHKVV	142	1492	A	1769	1	406	
143				17.07	•	, , , ,	
143							1
143					'		
143	!		l			,	
KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA NELCEVNRKGCTSGDPCLPYFCVQGCKLQQA SDF1ARQGTLIQVPSSAGEVECYYGISCGQSGL LENCMEMHCMDLPTDTSALVR 1494	143	1493	A	1789	1	447	
NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA SDFIARQGTILQVPSSAGEVECYXICSCGQSGL LENCMEMHCMDLPTDTSALVR 144	1				_		
SDFIARQGTLIQVPSSAGEVECYKICSCGQSĞL LENCMEMHCMDLPTIDTSALVR 1494 A 1814 J 404 PGRRFRPLSQAGTDSGS*VFPDSFPSAPAEPL PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI 145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 I 372 ERLVLTSEHCLVLTLFWFSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYJEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
LENCMEMHCMDLPTDTSALVR	ŀ						
PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI 145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHISREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPFKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNYYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRIFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI 145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHISREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPFKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNYYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRIFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	144	1494	A	1814]	404	PGRRFRPRLSOAGTDSGS*VFPDSFPSAPAEPL
IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI	i '						PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI]						KCNGEWVSQNDHVTQEGLDEATGLRVREVH
145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THILALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRIFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWHH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI,FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI,FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL 146	[l .
PEDARVLCRQLNCGTAL 146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRIFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	<u> </u>						
146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	(İ	İ	RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET 147	L		L				PEDARVLCRQLNCGTAL
CRKKTSVKRKTLEPLFDET 147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	146	1496	Α	1828	574	333	
147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	L					<u> </u>	CRKKTSVKRKTLEPLFDET
TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	147	1497	Α	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
YYPPLRSDTVRKFMVGSQMLAQAQRDLTPE Q 148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	}			!			VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQAITLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRI.FAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							TSVTYSMG*HGAPTGSEAGANWNH**LHAH
148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	L						Q
GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	148	1498	Α	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	ļ l				ļ		IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	i					:	GIEGRLTADQLNSATACIFAAEVAIKESERFN
RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							GIPALSVPVAEPIRHAEALMQQALTLKRSDET
149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	. [1		
149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
	149	1499	Α	1880	611	24	GLNSENALSNEAMERGWQCLRI.FAERLQDIP
VOVISGERFADI IVOGVARTTCCADODI VUD					Į		
VQ*ISOLLEARLI QO*AHI IOOADQRL*VD							VQVISGEEEARLIYQGVAHTTGGADQRLVVD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
			İ		-	YFADRNLGQENFDAAQKAAREVLRPVADEL RYHSWKEVRGASVTVQALQEIMMAQGMDE RITMEIWPVD
150	1500	Α	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS PESSILDGMIRQLQQQQDQRMGADQDTIPRG LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV EGVRQMHQNAPRSQIATERDLQAWKRRVVV PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA MEMQIKKQFQDTCKVQTKQYKALKNHQLEV TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI NEMMASQALRLDEAQEAECQALRLQLQQEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	Α	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL RQKIFKERALPDIENYMFENHDQLRQAATEC MCNMVLHKEVQERFLADGNDRLKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLLKDMDLLRGREAELKQRVEAFESYQLELK DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	À	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	Α	1939	1	423	TTHRLNYTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA
161	1511	Α	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucna		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
dence	İ	l	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
)]		sequence	/=possible nucleotide deletion, \=possible
	Į.			peptide	į	
	ļ	 		sequence		nucleotide insertion
į		1	ĺ		[ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
	1	l	1		1	PASDFSGALETDLKASLFDQPLSIICGDSDTLP
)	ł	ł	ł			RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
	1		l	Į		KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
L		L	<u> </u>			RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	Α	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
1	i	į			į	KFQGRWGTVCDDNFNIDHASVICRQLECGSA
ł	ł		ł	ł	1	VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
ł	1	ł	į		1	CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	A	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
	!					LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
	Ì				Ì	FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAOLFGG
107	1317	1 **	2012	207	1 37,	SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
	Ì					ENNWYFVVADSSKAGFTTIYKWERETGFYSH
	<u> </u>	i				
	L	<u> </u>			100	QSFTR
165	1515	Α	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
i !	(NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
	1		ļ		i .	DTHWRVAHERDELWRAQIVATTVMLERKLP
	Į.				ļ	RCLWPRSGICGREYGLGDRWILRVEDRQDLN
:		ĺ				RQRIQRYA
166	1516	A	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
ĺ	ĺ	ĺ				QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
	l					NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
1	•	ŀ				SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
ł	ł					WDMRNLATIFLAVVMALLSLHCLAAFKRLE
						HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
1						VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
						LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
!	i					GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
Ì		ĺ				HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	
107	1517	Λ.	2023	090	/1	AAASAASSLTVTLGRLASACSHSILRPSGPGA
						ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
						PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
						GRLFAVVHFASRQWKVTSEDLILIGNELDLA
						CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
						RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
						TTPQTVLRINSIEIAPCLL
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
						RLQGAARVFMPLQAQVKAKASKPLQMQIKA
,						PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
						KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	Ā	2049	1	945	ONLEDREVLNGVOTELLTSPRTKDTLSDMTR
						TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
				i		EDNSRSKREGLFHENECIVKINNVDLVDKTFA
						QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
						VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
						NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
						SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
						TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ
						SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
1.50						ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	Α	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
				1		VAKVEKTYDKTLENAVVADAVASKCSVLNE
						KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
		!		i		ESSSEESLGESKEQLGDDVTKPSSQKA
171	1521	Ā	2055	139	675	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS
			-		-	DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
						LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL
			1		L	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		RESCTITAYPCDSYQDYRNGKCVSCGTSQKE SCPLLGYYADNWKDHLRGKDPPMTKAFFDT AEESPFCMYHYFVDIITWNKNVR
172	1522	A	2056	3	361	LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK MEWKVKIRSDGTRYITKRPVRDRILKERALKI KEERSGLTTDDDTMSEMKMGRYWSKEERKQ HLVRGKEQRRRREFMMRIRLKCLKES
173	1523	Α	2060	1	387	GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA G
174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE RRLKEEEEARLKYEKEEMERLEIQRIEKEKW HRLEAKDLERRNEELEELYLLERCFPEAEKLK QETKLLSQWKHYIQCDGSPDPSVAQEMNT
175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV AESSTVGWLCALFRVTHVGVGATGHGVVCG RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ RLQFPYLEPGHELPATTLLAFLAAV
176	1526	A	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT GIHSVYTVYQGHEIMFHVSTMLPYSKENKQQ VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PCI
177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC CDGAWLAWACWVFGNDFPSPASAACSALLG CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ
178	1528	A	2104	2	409	ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD
179	1529	A	2111	1	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLACLG
180	1530	A	2116	3	366	TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV
181	1531	A	2117	2	386	YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI
182	1532	A	2123	1	493	RTKTDVYILNLAVADLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS
183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
184	1534	A	2145	3	538	RFSLIDDGAGPFCSAAYTTTGCRTPYL HELTVAAADRGQPPQSSVVPVTVTVLDVND NPPVFTRASYRVTVPEDTPVGAELLHVEASD ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW PKNNFNGSLVQASYQHEELRREVIMLACSFG NKHCHQQASTLISDWISSNRNRIPLNVRDIVY CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI DVIIHVARNPHGRDLAWKFFRDKWKILNTRI RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP LDEQNRDWQGLLENLHVELTLDEEDSEGPEK EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC AHDKSP
187	1537	Α	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDQPG OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP AVVVPYMMVLQENGYGVEEGIPTLLMAASS MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV AIVVSSLDW
190	1540	A	2179	64	399	MRLNONTLLLESFGXXRPYTSEHAPTYHQW MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR LSYVLFIQERDVHKGMFATNVTENVLNSSRV QEAIAEVAAELNPDGSAQQQSKAVNKVKKK AKRILQEMVATVSPAMIRLTGWVLLKLFNSF FWNIQIHKGQLEMVKAATETNLPLLFLPVHR SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS YTHSKGIMHRDVKPLNILCNSPRNKVILADW GLAEFYHPMRKYSVHVATRYYKSPEILLDYE YYDYSLDIWAVGVILLELLTLKLHVFEGGDN EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE LPVPMGARYIRINPQSWFDNGSICMRMEILGC PLPDPNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY ALYVHNCRLVFLLRKDFDQADTFRPAEFHW
196	1546	A	2256	1	396	KLDQQALAKVDGQPGKSITRQLQEMPVTIQG ISLKPS FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF
197	1547	A	2259	43	594	KALEESGALLESGTYDGNFYGTPKPPAEPSPF QPDPV QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL
198	1548	A	2275	3	404	HCFTESGRGKSWRLCAAILPL TCTTVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	Α	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM
201	1551	Α	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	A	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG MGANASTSPRCWDLSSGNKKWIIQVPILASIV ESRGGLLATGVGGMCACVPRNOPLTGT
207	1557	Α	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP NASNLDKVLTDIKADKDQANDGLSSALLILY LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG
211	1561	A	2431	1	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG
212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ
213	1563	A	2445		1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AMVDGRWLDCVSHHDQLFRDEYALYRPLQV LFSVYCQLECSKLIL
215	1565	A	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ HGPGRHGRRVCSSQDSMADVFVHLRTAWPT CSLISGQHGPGESVSYEDDDIPAPASLLHVNA AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSC TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSC TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH STVPGVRISSCTPDLTCAVSIHATVPGVRISSC TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSIHATVPGVRISSC TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH STVPGVCISSRTPDLTCAVSIHATVPGVRISSCT PDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTTDLTCAVSIHATVPGVRISSCT PDLTCAVSTHTTVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTTDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHLTVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSIHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVRISSCT PDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS
216	1566	A	2477	1	414	TPDLTCAVSHSTVPGLLTSVSQTSTG FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY
217	1567	A	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS
220	1570	A	2498		1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS LEGANAEIQGLKENLQNTNALNSQTQAFIKSS

SEQ ID	SEQ ID	Mct	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	peptide	מטת	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nuci-		Į.	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-				corresponding to last amino	
seq-	uence		09/496	correspondi	acid residue	M=Methionine, N=Asparagine, P=Proline,
uence	İ	ļ	914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
ì	i .	Í	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i	ļ		peptide		/=possible nucleotide deletion, \=possible
1	}	ŀ	ł	sequence	<u> </u>	nucleotide insertion
						FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
	l					VTAQTQKANGRLDQTDTQIQVFKSEMENVN
ł	j	}]	}	1	TLNAQIQVLNGHMKNASREIQTLKQGMKNA
1	[Į	1		SALTSQTQMLDSNLQKASAEIQRLRGDLENT
1	l			Ì		KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	Α	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV
1 221	13/1	^	2501	,	1 300	
	1	i	İ	Ì	ĺ	TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
	ĺ	l	į ·	İ		TCGEEASVLEILVYNSKIENRHEMLAVEPINE
ŀ					j	LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
	1	ļ	İ	1	•	AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT
	ł	L	L	ł		GVLFFFTN
222	1572	Α	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
	١.			1	[ARHVRDKEEEVDLVMQKVESLRQELRRTER
	j]	j	1	ļ	AKKELEVHTEALAAEASKDRKLREQSEHYSK
ı	[[QLENELEGLKQKQISYSPGVCSIEHQQEITKL
	!			1		KTDLEKKS
223	1573	A	2544	2	412	NDPAIISNFSAAVVHTIVNETLESMTSLEVTK
223	1373	Α	2344	4	412	
		l	1			MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
1		ĺ	[İ		EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
			ţ .			KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
		L]	LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
İ	!	ļ	1			CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
		ĺ			į	QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
		}		İ		TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
1	[1				VK .
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
223	1373	^	2505	/24	•	GLGEEEKEAGKKKKKQEEKEKGAVYSR
1	ł	Ì	l	}	}	VARICKNOMGGSQRVLEKHWTSFLKARLNC
1		ĺ	İ			SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
1						QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP
l		l	į			DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
		ł	Į į			TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
{						KTRVRYRLTAISVDHSAGPYH
226	1576	Ā	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
						TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
!						AGAACDRGMSLEACEAVTRKANRRTYTMG
1						VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
			i	i		RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
	13//	, A	23,3		1177	EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
	·		!			
1					}	ELWMKAMLDAALVQTEPVKRVDKITSENAP
		l	[TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
						KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
						PSEYESGSACPAQTVHYRPINLSSSENKIVNVS
		1				LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
			}			LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
						HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
						STHDKTLGPGAEEKRRSMRDDTMWQLYEW
			i l	ļ		QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
						MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
						QRGDVTIDRRHRAHHPKVK
228	1578	A	2583	3	330	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
220	1210	^	دورع	,	ا ،	
						HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
					•	KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD
		 _				PTMGIKPHLWWVAA
229	1579	A	2589	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
						ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
		L				GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met - hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GGTCLAGNLCTCPYGFVGPRCETMVCNRHC ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	Α	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI LYLLYLLRNAFFLR
231	1581	Α	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA GRARRTPTCEPATPLCCRRDHYVNFQELGW RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS FHSAVFSLLKANNPWPGRTSWCVPTARRPLS LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ NSIPYWERIT
234	1584	A	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTĞVASTRHEMGTLDKHKELEDLV AKFLNVEAAMVFGMGFATNSMNIPALVGKG CLILRDEVNHTSLVLGARLLGATIGIFKHNYA QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS FGASGGYIAGRKARILSPPACLVPNTGSHSLH RLTRDLQMNEAMVALVTDRLQGWNSGEGN WDRADKFGDLVDYLRVHSHSAVYASSMSPPI AEQIIRSLKLIMGLDGTTQ
236	1586	A	2621	1	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP A
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV
238	1588	A	2631	1	1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN
239	1589	A	2636	1	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI
240	1590	Α	2639	389	3	ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD LWQNFAREIEEHVFTLYSKNIKKYKTCIRSKV ANI.KNPRNSHLQQNLI.SGTTSPREFAEMTVM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1591	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	Α	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	Α	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1		MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	Α	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	A	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	Α	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
						PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI GAGDCL
252	1602	A	2697	421		PQKSHSGAYQCFATRKAQTAQDFAIIALEDG TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT VTWALDDEPIVRDGSHRTNQYTMSDGTTISH MNVTGPQIRDGGVYRCTARNLVGSAEYQARI NVRGPPSIRAMRNIT
253	1603	Α	2698	65	401	ACCQWRRTLIPAKSTTVSCTISTPHHPFRGSYS FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA DKETLENMMQRHEEEAHEKGKILSEQKAMIN AMDSKIRSLEQRIVELSEANKLAANSSLFTQR NMKAQEEMISELRQQKFYLETQAGKLEAQN RKLEEQLEKISHQDHSDKNRLLELETRLREVS LEHEEQKLELKRQLTELQLSLQERESQLTALQ AARAALESQLRQAKTELEETTAEAEEEIQALT VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	Α	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAIQII.LQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ PYEAARMFFEGLR
257	1607	A	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD FNPSFSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL
258	1608		2709		1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR EIYLQGCFKPLVSISPNDSLFFAVYTLIKNRIH

RPSFLYRTIQDLGIGTFRDLAVVLETAPILTA DIFVDRVSALAVVNECGTHPQDERLGIGW GLGEPGSERLFPAITSR GPEFPGSAKLVFLDLSYNNLTQLGAGAFRS/ GRLVKLSLANNNLVGVHEDAFETLESLQVL LNDNNLRSLSVAALAALALALRSLRLDGNPW CDCDFAHLFSWIQENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV ESRRISLRACRRPASRV	In mucleotide USSN Ogregory Incation Ogregory Incation Ogregory Ogregory Incation Ogregory
Sequence USSN O9/496 Corresponding February Corresponding C	Sequence USSN Ogardo Corresponding Ogardo Oga
sequence uence 99496 oprresponding to first amino acid residue of peptide residue of peptide sequence peptide sequence	uence 09/496 correspondi 10 last amino 20 14 15 15 15 15 15 15 15
United Services	914 ng to first amino acid residue of peptide residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Tryptophan, Y-Stop codon, /-possible nucleotide deletion, -possible nucleotide insertion RI-PVLDPVSGNVLHILTHERLIKFILHIFGSLLI RPSFLYRTIQDLGIGTFRDLAVVLETAPILITAL
amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence sequence peptide sequence peptide sequence sequence peptide seque	amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence per per per per per per per per per pe
residue of peplide sequence	residue of peptide sequence Y=Tyrosine, X=Unknown, *=Siop codon, /=possible nucleotide deletion, k=possible nucleotide insertion RLPVLDPVSGNVLHILTHKRLLKFLHFGSLLL RFSL-YRT[QDLG]GTFRDLAVVLETAPILTAL DIFVDRRV\$ALAVVNECGTHPQDERLGLGW GLGEPGSERLFPAAITSR GRLVKLSLANNILGVHEDAFETLESLQVLL LNDNNLRSLSVALALALPALRSLRIDGNPWL LNDNNLRSLSVALALALPALRSLRIDGNPWL CDCDFAHLFSWIQENASKLPKGLDEIQCSLPN ESRRISLRACRPASRV CDCDFAHLFSWIQENASKLPKGLDEIQCSLPN ESRRISLRACRPASRV CDCDFAHLFSWIQENASKLPKGLDEIQCSLPN ESRRISLRACRPASRV ARSSGNRTFQNRNFHLRLQRSQKKGPFA RLISPLVNLPQSPGGLEFQVQAT RAMLKCLREGQPPSYNWIRLDGFLPSGVKKGPFA RLISPLVNLPQSPGGLEFQVQAT RAMLKCLREGQPPSYNWIRLDGFLPSGVKNDGPHLTV CDGDTLGPPLTTELSGVYRHDTNEFSSRDSH DITVDVLDPPEDSGKQVDL LAAVETTVLVLIFAVSLLGNVCALVLVARRR RRGATACLVLNLFCADLLFISAIPLVLAVRWT EAWLLGPVACHLLFVVMTLSGSVTLLAAV SLER LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V
peptide	peptide
RIPVLDPVSGNVLHILTHKRLLKFLHIFGSLI RRSFLYRIQDLGIGITRDLAVVETAPILTA DIFVDRRYSALAVNECGTHPQDERIGLGU GLGEPGSERLFPAAITSR GREFPGSAKLVFLDLSYNNITQLGAGAFRS/ GRLVKLSLANNNLVGVHEDAFETLESLQVL LNDNNLRSLSVAALAALPALRSLRLDGNPW CDCDFAHLFSWOGNASKLPKGLDEIQCSLP ESRRISLRACRRPASRV 263	RLPVLDPVSGNVLHITHKRLLKFLHIFGSLI RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL DIFVDRRVSALAVVNECGTHPQDERIGLGW GLGEPGSERLFPAJITSR GPEPFGSAKLVFLDLSYNNLTQLGAGAFRSA GPEPFGSAKLVFLDLSYNNLTQLGAGAFRSA GRLVKLSLANNNLVGVHEDAFETLESLQVLE LNDNNLRSLSVAALAALPALRSLRLDGNFWI CDCDFAHLFSWIGENSKLPKGLDEIQCSLPW ESRRISLRACRRPASRV BISH DEVELOPENSKLPKGLDEIQCSLPW ESRRISLRACRRPASRV ASSSQNNRTFQNNRNFLRLQRDSQKKGQFA RLISPLVNLPQSPGGLEFQVGVA ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA RLISPLVNLPQSPGGLEFQVGVA ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA RLISPLVNLPQSPGGLEFQVAVA ARSSQLANTFREPSPSVKGDHRILV LAAVETTVLVLIFAVSLLGNVCALVLVARRR RRGATACLVLNLFCADLLFISAPILVLAVRWT EAWLLGPVACHLLFYWMTLSGSVTILTLAAV SLER LGVLPQAFHIFQDFLTGDLFWPELLGLP V
RPSFLYRTIQDLGIGTFRDLAVVLETAPILTA DIFVDRVSALAVVNECOTHPQDERLGIGW GLGEPGSERLFPAITSR GREPGSERLFPAITSR GPEFPGSAKLVFLDLSYNNLTQLGAGAFRS/ GRLVKLSLANNNLVGVHEDAFETLESLQVL LNDNNLRSLSVAALAALPLALRSLRLDGNPW CDCDFAHLFSWIQENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV CDCDFAHLFSWIQENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV CDCDFAHLFSWIQENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV ARSSQNNRTFQNNRNFLRLQRDSQKKGQF RLISPLVNLPQSPGGLEFQVQAT ARSSQNNRTFQNNRNFLRLQRDSQKKGQF RLISPLVNLPQSPGGLEFQVQAT DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSI DTVDVLDPPEDSGKQVDL DTVDVLDPPEDSGKQVDL LAAVETTVLVLIFAVSLLGNVCALVLVARRI RRGATACLVLNLFCADLLFISAPLVLAVRW EAWILGPVACHILFYVMTLSGSVTILTLAA' SLER LGGYLPLQAFHIFQDPLTGDLPWPELILGIL V	RPSFLYRTIQDLGIGTFRDLAVVLETAPLITAL
DIFVDRNYSALAVVNEGGTHPQDERLGLGW GLGEPGSERLFPAAITSR GLGVEPGSERLFPAAITSR GLGVEPGSERLFPAAITSR GLGVEPGSERLFPAAITSR GRLVKLSLANNNLVGVHEDAFETLESLQVL LNDNNLRSLSVAALAALPALRSLRLDGNPW CDCDFAHLFSWIQENASKLFKGLDEIQCSLP ESRRISLRACRRPASRV PARISGVDPPVRKATKGGENCSFEDNKNWQ LWGLNGAFFFKEFWGGRNNHAKGFRTTW ARSSSQNNRTFQNNRNFFLQPDSQKKGOF RLISPLVNLPQSPGGLEFQYQAT CSG STEDNKNWQ LWGLNGAFFFKEFWGGRNNHAKGFRTTW ARSSSQNNRTFGNNRNFLQRDSQKKGOF RLISPLVNLPQSDGKGOF RLISPLVNLPQSDGKGOF RLISPLVNLPQSDGKGOF RLISPLVNLPQSDGKGOF RLISPLVNLPQSDGKGVDL LAVETTVLVLIFAVSLLGNVCALVLVARRI RRGATACLVLNLFCADLLFISAFLVLAVRW EAWLLGPVACHLLFYVMTLSGSVTILTLAAT SLER LQGYLPLQDAFHIFQDPLTGDLPWFELIGLI V V V SRIPPL LYFQTHKGLKDSGKQVDL V V ASRLEKQNSTPESDYDNTFNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKOFTTHTPL LYFQTHKGLKDSSEVTCLGISQCWRKGF AVLLLLLLSLALGLVLAALGLFVHRRDSPL VQASGGPLACFGLVCLGLVCLSVLLFFQGW ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE LPLSWAE LPLSWAE LPLSWAE LAVENTY LYFQTHKGLKDSGLVLLFFQCW ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE LP	DIPVDRRY \$\text{3} ala \text{VNECGTHPQDERLGLGW} \\ GLOUD
GLGEPGSEERLFPAAITSR	GLGPPGSERLFPAAITSR
262	1612
GRLVKLSLANNNLVGVHEDAFETLESLQVL LNDNNLRSLSVAALAALPALRSLRLDGNPW CDCDFAHLFSWIQENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV	GRLVKLSLANNNLVGVHEDAFETLESLQVLE LNDNNLRSLSVAALAALPALRSLRLDGNPWI CDCDPAHLPSWIQENASKLPKGLDEIQCSLPM ESRRISLRACRRPASRV 1613
LNDNNLRSLSVAALAALPALRSLRLDGNPW CDCDFAHLFSWIGENASKLPKGLDEIQCSLP ESRRISLRACRRPASRV 263	LNDNNLRSLSVAALAALPALRSLRLDGNPWL CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
CDCDFAHLFSWIQENASKLPKGLDEIQCSLP	CDCDFAHLFSWIQENASKLPKGLDEIQCSLPNESRRISLRACRPASRV
ESRRISLRACRRPASRV	ESRRISLRACRRPASRV
263	1613
263	1613
LWGLNGNFNFFKEPWGGRNNHAKGFRTTW ARSSSQNNRTFQNNRNFLRLQRDSQKKGQF RLISPLVNLPQSPGGLEFQYQAT RAMLKCLREGQPPFSYNWTRLDGFLPSGVR DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSI DTVDVLDPPEDSGKQVDL DTVDVLDPPEDSGKQVDL DTVDVLDPPEDSGKQVDL LAVETTVLVLIFAVSLLGNVCALVLVARRI RRGATACLVLNLFCADLLFISAIPLVLAVRW EAWLLGPVACHLLFYVMTLSGSVTILTLAA' SLER AFREVGGYWGLLCEHLYAIPSKTSEGNWTA LQGYLPLQDAFHIFQDPLTGDLPWPELILGLI V	LWGLNGNFNFFKEPWGGRNNHAKGFRTTW ARSSQNNRTFQNNRNFLRLQRDSQKKGQFA RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV DGDTLGFFPPLTTEHSGIYVRHDTNEFSSRDSH DTVDVLDPPEDSGKQVDL A 2752 2 388
ARSSQNNRTFQNNRNFLRLQRDSQKKGQF RLISPLVNLPQSPGGLEFQYQAT	ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
RLISPLVNLPQSPGGLEFQYQAT	RLISPLVNLPQSPGGLEFQYQAT
264	1614
DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSI DTVDVLDPPEDSGKQVDL 265	DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH DTVDVLDPPEDSGKQVDL
DTVDVLDPPEDSGKQVDL	DTVDVLDPPEDSGKQVDL
265	1615
LAAVETTVLVLIFAVSLLGNVCALVLVARRI RRGATACLVLNLFCADLLFISAIPLVLAVRW EAWLLGPVACHLLFYVMTLSGSVTILTLAAY SLER 266 1616 A 2755 192 I AFREVGGYWGLLCEHLYAIPSKTSEGNWTA LQGYLPLQDAFHIFQDPLTGDLPWPELILGLI V 267 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGPLACFGLVCLGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLRREVSEHGRHLQQYFNLFVMYAY LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDIK KHALERFEQEAVAVDSNLRVRELQRKVDGI	LAAVETTVLVLIFAVSLLGNVCALVLVARRR RRGATACLVLNLFCADLLFISAIPLVLAVRWT EAWLLGPVACHLLFYVMTLSGSVTILTLAAV SLER 1616 A 2755 192 1 AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHI LIAVLNLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVIVQLPTESIQKNQEDKLIKMVPRKQRE FSGSDRGKLPGSEGKNQEDSKLIKMVPRKQRE FSGSDRGKLPGSEGKNQEDSKLIKMVPRKQRE FRKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
RRGATACLVLNLFCADLLFISAIPLVLAVRW EAWLLGPVACHLLFYVMTLSGSVTILTLAAT SLER 266 1616 A 2755 192 I AFREVGGYWGLLCEHLYAIPSKTSEGNWTA LQGYLPLQDAFHIFQDPLTGDLPWPELILGLI V 267 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAT LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDIG KHALERFEQEAVAVDSNLRVRELQRKVDGI	RRGATACLVLNLFCADLLFISAIPLVLAVRWT EAWLLGPVACHLLFYVMTLSGSVTILTLAAV SLER 1616 A 2755 192 1 AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHI LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNRSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
EAWLLGPVACHLLFYVMTLSGSVTILTLAA' SLER 266 1616 A 2755 192 1 AFREVGGYWGLLCEHLYAIPSKTSEGNWTA LQGYLPLQDAFHIFQDPLTGDLPWPELILGLI V 267 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF 268 1618 A 2762 1 405 IACTFCGQDEWSPERSTRCFRRSRFLAWGE AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAI LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	EAWLLGPVACHLLFYVMTLSGSVTILTLAAV SLER
SLER	SLER
266	1616 A 2755 192 1 AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
LQGYLPLQDAFHIFQDPLTGDLPWPELILGLI V 267 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF 268 1618 A 2762 1 405 IACTFCGQDEWSPERSTRCFRRRSRFLAWGE AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGVLLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAI LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V 1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGVLCSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
V ASRLEKQNSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF	V
267 1617 A 2760 434 714 ASRLEKONSTPESDYDNTPNDMEPDGMGYN HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269	1617 A 2760 434 714 ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF AVLLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
HRTSVPGEGLPRARDLAGLGQQKQFTTHTPI LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF 268 1618 A 2762 1 405 IACTFCGQDEWSPERSTRCFRRRSRFLAWGE AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAL LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF 1618 A 2762 1 405 IACTFCGQDEWSPERSTRCFRRRSRFLAWGEP AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 1 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
LYFQTHKGLKDSSIRSEVTCLGISQCWRKGF	LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF 1618 A 2762 I 405 IACTFCGQDEWSPERSTRCFRRRSRFLAWGEP AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
268	A 2762 1 405 IACTFCGQDEWSPERSTRCFRRSSFLAWGEP
AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAI LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	AVLLLLLLSALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
VQASGGPLACFGLVCLGLVCLSVLLFPGQPS ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAI LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGII	VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVES LPLSWAE 269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYAI LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDH KHALERFEQEAVAVDSNLRVRELQRKVDGH	ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
LPLSWAE	LPLSWAE LPLSWAE LPLSWAE LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY LOSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP LOSQUACHTRETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP LOSQUACHTRETEQLRSQLQTTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
269 1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDH LIAVLNLLRREVSEHGRHLQQYFNLFVMYA LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDH KHALERFEQEAVAVDSNLRVRELQRKVDGH	1619 A 2772 3 243 TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHIL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
LIAVLNLLRREVSEHGRHLQQYFNLFVMYAL LSKNLSFSEFCFDVSY 270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDF KHALERFEQEAVAVDSNLRVRELQRKVDGI	LSKNLSFSEFCFDVSY 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KERVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
270 1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQ VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	1620 A 2789 I 486 ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
VEQIAKAEETHSSLSQELQARLQTVTREKEE LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
LQLSIERGKVLQNKQAEICQLEEKLEIANEDI KHALERFEQEAVAVDSNLRVRELQRKVDGI	LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
KHALERFEQEAVAVDSNLRVRELQRKVDGI	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
KHALERFEQEAVAVDSNLRVRELQRKVDGI	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP 1621 A 2795 1 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
271 1621 A 2795 I 568 KEKRVTVQLPTESIQKNQEDKLKMVPRKQR	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
VKGIOEKOVFSNTKDSKOEITONKSFFSSVKI	
	SQRDDGKGALNIVEFLRKREELHQILSTVKQP
	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
	GYRKVVSNNCTDGVREOYTAKPOKCPGKAP
	RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
NORTH INDUSTRIACIONIN'S LIMIVOLEBOID VODEL INDUSTRIACIONIS ESE EDITORIS I	VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
YQNXGIXRXTVQVDNSLGS	
	DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
AVSRKTELEEYH	
274 1624 A 2805 168 320 ILWLYFETGTWVYPVFAKLSLLGLAALFSLR	
IFIARNGVVGETLTHCKRV	624 A 2805 168 320 ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l		l	{	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			}	peptide		/=possible nucleotide deletion, \=possible
055	1.635	 	1	sequence		nucleotide insertion
275	1625	Α	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
276	1626	A	2813	41	266	MGKIIFQ AGRSLHGAGDRAWVGISPTDWSPKVVELCK
270	1020	Α .	2013	(41	200	KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
		ł				QVGPVRRNGEAGPG
277	1627	A	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
		1	-017	1	1 .20	LFISYLHTPKHKQHEVLQAMGSILGITGEEME
·		1		}		PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
		ì				GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
		İ			1	LPPHNSPGKIK
278	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
				İ		VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
	<u> </u>	<u>L</u>	<u> </u>			PVDQNPRLV
279	1629	Α	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
	1					TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
					1	CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
200	1620	 	2026	200	l	VPSQRHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
		[ĺ	CGQYWPLEKDSRIRFGFI.TVTNI.TGAVGEPG VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
201	1031	^	2021	01	301	QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
1			1			NTTNMDEVPRPQALSGSSVVWVSGCVASRS
						VILSLTSG
282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
	1					TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
						YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
		1				YLKTLPPYYL
283	1633	Α	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
İ	1	i	1			MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
i	ļ					PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
284	1634	_	2027	2	204	SSEESAP
284	1034	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ DYGLVVRGCLDLRYLAMRORNNLLCNGLSL
						KSLAETVLNFPLDKSLLLRCSNWDAETLTED
		i				QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
ĺ		1				KR
285	1635	A	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
j						VCDRVSEDGINRQQAQEWCIKHGFELVELSP
	_	}				EELPEEDGKCLCVRRKYGTYI
286	1636	Α	2845	197	278	TAEDVLTVAYEHGVNLFDTAEVYAAGK
287	1637	A	2851	2	427	FVAEVRREWAKYMEVHEKASFTNSELHRAM
ļ •	1	1				NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
1			1 1			AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
						QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
200	1/22	 	0050		460	KVYLEQNLAAQDRVLCALT
288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
		i				LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
						TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIOVYAEGVELMEPLL
						EPGOELAETALHLAVRTADOTSLHLVE
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
]	l	ا """	-		DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
			1			DSGLWRMHLMEGELPASMSGSCGACINGKL
	1		1 1			YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
						ITDFEGQPPTPRDKLSCWVYKDRLIYFG
290	1640	1	2868	1	378	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
	1040	Α	2000 1	• !	3.0	
ļ	1040	A	2808	•		SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF PDCASCLQAQDPLCGWCVLQGRCTRKGQCG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496 914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
)	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Jaquanaa	/=possible nucleotide deletion, \=possible
	İ	l		sequence	1	nucleotide insertion
				7		RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR
				L		QE
291	1641	Α	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
1	•	Ì	į	<u> </u>		PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
		ŀ	İ		•	GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
		Ì				AKVINAENAAHKSEKFRAMATRTRQEYLKD
292	1642	A	2877	3	188	LA DIVERDIDA TEOGRAPHICA MARIA
292	1042	l A	28//	3	198	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	1	427	REKEEEVEEEDKVVKETEKEAEQEKEEDSL
2,53	1043	<u> </u>	2070	1	427	GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
· .		l				LSPEKLTAENRYYCESCASLQDAEKVVELSQ
		{			Ì	GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
		1				LRLPLAGGRGQAYDL
294	1644	A	2879	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI
		ļ				IIVFVTGGVLG
295	1645	Α	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
		ĺ			ĺ	NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF
	ŀ				ļ	FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
-204	1646		2002	200	262	LSESISQ
296	1646	Α	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
297	1647	Ā	2893	8	424	RLQEFSQKMDQVRGHWPVST SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
251	1047	Α .	2093	0	424	KLYSTMGRFLRDRKNPACREMAVVLLANLA
						QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT
'				İ		QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
						LLALAKVDDNHSEF
298	1648	Α	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
						SGLLNASAQVNL
299	1649	Α	2898	1	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL
						GYFQAYNVLILTMQASLPKVLRFCACAGMIY
)			LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
						VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
						LQEF
300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
					- 1,2	TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
			1			SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ
						LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
						FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
			1			EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
302	1650	A	2909	2	412	KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	A	2707	4	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
						CERHFOKVLNRSLFTGLRSITHFGRPPFEPFFN
				'		SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
						LVNRQDRAHFM
303	1653	Α	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
				•	- -	VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
					<u> </u>	NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	Α	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
						DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
						PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE
100	1767		204:			LHAREWP
306	1656	A	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
	i	!			İ	EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA
<u>. </u>			Ll			LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SKTKLSFKVSSSA
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT CTAENAVGRARRRVHLTILVLPVFTTLPGDRS LRLGDRLWLR
308	1658	Α	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM DSSLPEEEEDEDKEAINGSGNAENRERHSESS DWMKTVPSYNQTNSSMDFRNYMMRDETLEP LPKNWEMAYTDTGMIYFIDHNIKTTTWLDP RLCKKAKAPEDC
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI TGFVQLSISVTALTAILKYGQVLMHSHVVIIW LFLAVYAVATIMFCF
311	1661	A	2963	3	465	MKPOMPGLGAPNGYGPGRGRAGVPGGPERR PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK PQKPGLRGTLKPQKSGHGHENGPWPGPCNA RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ KI
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ
314	1664	A	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM
315	1665	Ā	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE
317	1667	A	2981	3		VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARI.I.RIDIANTI.REQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion STLALSHSAQVLASASGRSSTTAHCQIRVWD
320	1670	Ā	3000	693	322	VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL GDHDGRTLALWGTGHL IDESTGLIITVNYLDYETKTSYMMNVSATDQA
						PPFNQGFCSVYITLLNELDEAVQFSNASYEAA ILENLALGTEIVRVQAYSIDNLNQITYRFDAY TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF LFPWLFLQVEVIKKAYMQGEVEFEDGENGK DGAASPRNVGHNIYILAHOLARH
323	1673	Ā	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND ERVFGKRGF
324	1674	Ā	3020	523	797	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	Ā	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ EGDLVEVVLSASATFEDFQIRPHALTVHSYRA PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK RC
328	1678	A	3030	13	569	ITRPTISCORPGPGLAAGMLPYTVNFKVSART LTGALNAHNKAAVDWGWQGLIAYGCHSLV VVIDSITAQTLQVLEKHKADVVKVKWAREN YHHNIGSPYCLRLASADVNGKIIVWDVAAGV AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI HPPNYIVLWNADTGTKLWKKSYADNILSFSF D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL HRMAEKVGADITVLREREVDYDSDMPRKITE VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL LGVLTQGELDNGRGRARLNLFRHLHEIQSGR TSSISFEILGFNSKGEVHGINGTQWGQTLRMG W
330	1680	Α.	3040	3	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ VVLTMTNMGPVDTATYYCAQFARGARGSN WFDPWGQ
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK GENRKTLISGMIDEPHAIVVDPLRGTMYWSD WGNHPKIETAAMDGTLRETLVQDNIQWPTG LAVDYHNERLYWADAKLSVIGSIRLNGTDPI VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG KRLDNGTCVPVPSPTPPDAPRPGTCNLQCFN GGSCFLNARRQPKCRCQPRYTGDKCELDQC

SEQ ID NO: of NO: of NO: of nucleotide peptide eotide sequence USSN D9/496 914 nucleotide sequence USSN D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 914 or sequence D9/496 O	FTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF IVLKAV VQRLQ REGIRY CSLAW
nucleotide sequence Deptide sequence	EFTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQP RQISLVG TLPWGT MIPTAD EPREVLF EPREVLF EPREVLF VVQRLQ REGIRY VCSLAW
eotide sequence ue	EFTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQP RQISLVG TLPWGT MIPTAD EPREVLF EPREVLF EPREVLF VVQRLQ REGIRY VCSLAW
sequence Sequence 914	FTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG ILPWGT MIPTAD EPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 915 peptide sequence 916 917 Peptide sequence 918 919 919 919 919 919 919 91	FTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG ILPWGT MIPTAD EPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
amino acid residue of peptide sequence	GFTGPKC CRCLPG CRCLPG CRCLPG VVNKQS IGGSCT FSQQQP RQISLVG ILPWGT MIPTAD SPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon /=possible nucleotide deletion, \=possible nucleotide insertion WEHCRNGGTCAASPSGMPTCRCPTC TQQVCAGYCANNSTCTVNQGNQPQ FLGDRCQYRQCSGYCENFGTCQMA CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTU LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	GFTGPKC CRCLPG CRCLPG CRCLPG VVNKQS IGGSCT FSQQQP RQISLVG ILPWGT MIPTAD SPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion WEHCRNGGTCAASPSGMPTCRCPTC TQQVCAGYCANNSTCTVNQGNQPQ FLGDRCQYRQCSGYCENFGTCQMA CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHLASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDW LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	EFTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF IVLKAV IVQRLQ REGIRY IVCSLAW
sequence nucleotide insertion WEHCRNGGTCAASPSGMPTCRCPTC TQQVCAGYCANNSTCTVNQGNQPQ FLGDRCQYRQCSGYCENFGTCQMA CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	GFTGPKC CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF IVLKAV IVQRLQ REGIRY ICSLAW
WEHCRNGGTCAASPSGMPTCRCPTC TQQVCAGYCANNSTCTVNQGNQPQ FLGDRCQYRQCSGYCENFGTCQMA/ CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF IVLKAV IVQRLQ REGIRY ICSLAW
TQQVCAGYCANNSTCTVNQGNQPQ FLGDRCQYRQCSGYCENFGTCQMA CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	CRCLPG ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF IVLKAV IVQRLQ REGIRY ICSLAW
FLGDRCQYRQCSGYCENFGTCQMAN CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSQ EDRTRAHADHVQQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	ADGSRQ VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
CRCTAYFEGSRCEVNKCSRCLEGAC GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	VVNKQS IGGSCT FSQQQP RQISLVG TLPWGT MIPTAD EPREVLF EVLKAV VVQRLQ REGIRY VCSLAW
GDVTCNCTDGRVAPSCLTCVGHCSN MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVO QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	GGSCT FSQQP RQISLVG TLPWGT MIPTAD PREVLF VVLKAV VVQRLQ REGIRY VCSLAW
MNSKMMPECQCPPHMTGPRCEEHV GHIASILIP 332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVO QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	RQISLVG TLPWGT MIPTAD PREVLF VVLKAV VVQRLQ REGIRY VCSLAW
332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPMI AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVO QFGEWSDQPRITKDVICFHAEDFTU LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTI PASE	RQISLVG TLPWGT MIPTAD PREVLF VVLKAV VVQRLQ REGIRY VCSLAW
332 1682 A 3045 3 952 TTTISNFHTQVNRTYCCGTYRAGPM AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	TLPWGT MIPTAD PREVLF VLKAV VQRLQ REGIRY CSLAW
AVDEEVGDYFPEFLDMLEESPFLKM LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	TLPWGT MIPTAD PREVLF VLKAV VQRLQ REGIRY CSLAW
LSSLRLQCRSQSDDGPIMWVRPGEQI MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMI ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	MIPTAD PREVLF VLKAV VQRLQ REGIRY CSLAW
MPKSPFKRRSMNEIKNLQYLPRTSE EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	PREVLF VLKAV VQRLQ REGIRY CSLAW
EDRTRAHADHVGQGFDWQSTAAVG QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	VLKAV VQRLQ REGIRY 'CSLAW
QFGEWSDQPRITKDVICFHAEDFTDV LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	VQRLQ REGIRY CSLAW
LDLHEPPVSQCVQWVDEAKLNQMR ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	REGIRY CSLAW
ARIQLCDNDIYFIPRNVIHQFKTVSAV HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	CSLAW
HIRLKQYHPVVEATQNTESNSNMDC ELEVDSQCVRIKTESEEACTEIQLLTT PASE	
ELEVDSQCVRIKTESEEACTEIQLLTT PASE	GLTGKR
PASE	
333 1683 A 3046 497 167 SACSTGPELPGRATRSLTRPANOKGO	DGDRL.
YYDGCAMIAMNGSVFAQGSQFSLDI	
ATLDLEDVRSYRAEISSRNLAVSAPV	
CSSKTWKVAPFVRAWWRP	
334 1684 A 3053 37 276 VITDLEEQLNQLTEDNAELNNQNFYI	SKOLD
EASGANDEIVQLRSEVDHLRREITER	EMOLTS
QKQVRRVNKVVRSLEDF	
335 1685 A 3054 2 846 WDAWGDWSDCSRTCGGGASYSLRR	CLTGR
NCEGONIRYKTCSNHDCPPDAEDFRA	
YNDVQYQGHYYEWLPRYNDPAAPO	
AQGONLVVELAPKVLDGTRCNTDSL	DMCISG
ICQAVGCDRQLGSNAKEDNCGVCAC	DGSTC
	SRSVRI
TVKGPAHLFIESKTLQGSKGEHSFNS	PGVFVV
ENTTVEFQRGSERQTFKIPGPLMADF	
TAAKDSVVQFFFYQPISHQWRQTDFI	PCTVT
CGGG	
336 1686 A 3058 54 347 VVGKQEAGAHSDSCCLLHTPPRLTPA	
LRNSRIVSQKDDVHVCIMCLRAIMN	
AWDWRLGSPACPHWGLHKLPRLWI	PLSLYP
VLCWGT	
337 1687 A 3059 2 709 ILTSLVELTRFETLTPRFSATVPPCWV	
QQQRRHPQHLHQQHHGDAAQHTRT	
DSNSWDEHVFELVLPKACMVGHVD	
SNITNIPQIQVTLLKNKAPGLGKVNG	
LEDHKEDILCGPVWLASGLDLSGHA	
SPKLVKGMAGGKYRSFLIHVKAVNE	
CNGGMRPVVRLPSLKHQSNKGYSLA	SLLAK
VAAGKEKSSNVKNENTSGTRK	
338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLE	
DKVYGVADSCTSLLSGRNRCKLGLL	
SDVNPRNTFGQLFCGSLDLFGILCVG	LYRIIDE
EELNP	
339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVA	FQNYV
PSSVTTMLSWV	
340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLS	
LLNCHNKSAIDLAPTPQLKERLAYEF	KGHSLL
QAAREADVTRIKKHLSLEMVNFKHP	
LHCAAASPYPKRKQICELLLRKGANI	
FLTPLHVASEKAHNDVVEVVVKHEA	KVNAL
DNLGQTSLHRAAYCGHLQTCRLLLS	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	под	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}	ļ	ļ	ļ	peptide	1 .	/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
		<u> </u>	1			IISLQGFTALQMGNENVQQLLQEGISLGNSEA
				ĺ	_	DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE
	ĺ		ĺ		{	GRQSTPLHFAAGYNRVSVVEYLLQHGADVH
	1					AKDKGGLVPLHNACSYGHYEVAELLVKHGA
		1				VVNVADLWKFTPLHEAAAKGKYEICKLLLQ
	l]	l	ļ		HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
		İ	1		!	GDAALLDAAKKGCLARVKKLSSPDNVNCRD
						TQGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA
	[ĺ		ĺ	ĺ	NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI
						RTVKFLRSATIPVVELMDVQGERLDMEVGFD
						NRQAAFDMVCTMLEKRVRHKILYLG\$KDDT
	ł	ł		}	ł	RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI
	1600	.	0000	160		HLRMQLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE
						IDDNRVTAEEVDILLREGEKLAPVMAKTRILR
						AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE
,	1	1			ĺ	RDGLDGFITITGGKLMTYRLMAEWATDAVC
343	1693	A	3075	250	1	RKLGNTRPCTTADLALPGSQEPAKVP LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS
343	1073	Α.	3075	230	1	LGASRAQVLWFVILPGALPEILTGLRIGLGVG
		į			1	WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLOSLQVAAISTFCCLLIGYPLAWAV
	1000		30,0	ا آ		AHSKPSTRNILLLL
345	1695	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC
					-	VINQAAATGGDARQLVGYLVSQSGLPLDTSA
		İ				LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL
		ŀ				DRKALPLPELKAQAPGRAPKAGSETIIAAAFS
		l				SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	Α	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI
						ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI
	ļ		ł			ALSQSRSQKELEDVLYSNLEELTRMAKMVSD
				•		MLFLAQADNNQLIPEKKMLNLAHEVGKVFD
347	1697		3084	_	340	QFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE
	1		1		.	DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT
	<u> </u>		1			OPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL
J-10 .	1076	^	3000	143	.0	YNIAAYPHLKGDDLAEQAQALSNRAYEEAA
	ļ	ļ	1			QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG
		1				PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI
	!					AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK
	1	1	l			ALPNVPWVDHTRVAAFGFRFGANVAVRLAY
	1	ľ	:	-		LESPRLKAVACLGPVVHTLLSGLKCOOVPE
•			ł			MYLDVLASRLGMHDASTKSSTRENH
349	1699	Α	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV
1	ŀ	ļ				GFGVAMSQALGPFSLRAGVASSTLGIAQVCG
		_				SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR
		ŀ	}		[ANTPDSDITEKTEDSSVPETPDNERKASISYFK
						NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN
	<u> </u>					KDTVIIVSEPSEDEESQGLPTMARRNDDISELE
		L				DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG
						MFGIVVMVIETELSWGAYYKAPLYSLALKCL
						ISLFTIILLGLTIVYHAREIQLFMANYGADDWR
] .			SALTYEPIFLILLEALRGVIHATPCRVSLSLWD
	l		L			GLDLP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence]	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l	i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	ļ	/=possible nucleotide deletion, \=possible
		-		sequence	ļ	nucleotide insertion
352	1702	Α	3110	341	2	AQLAEVCPPQTLLTTNTSSISITALAAEIKNPER
1	i					VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
l		1	Ì		ł	QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
353	1703	 	2111		100	YSEAWRALEEQVAAPEVI
333	1 /03	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
354	1704	A	3116	367	225	GYGNTMVGIAVGIQFLATVLTRGYAGRLA
334	1 /04	A	3116	307	223	WQLFFILNGTFLNIGETDTESCVNGWVYDRSS FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
333	1703	A	3117	101) 33	EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
		ļ			ŀ	ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
1			[RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
1						LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
						SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
i		ĺ	[DEPTN\HLGHGRAIEALGPCLOTISGVGVILVS
						HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	Ā	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG
330	****	ı	3.2.	137		AASREHARWOGTGLAPGTRVAVAPTCVOGL
1	}	ļ				PQERSVCRPFFSSRWREGPVWALGAGAHGKP
						RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGOMLEVO
1 33,	''''	'	3.2.	1277	1227	PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
		[ļ			EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
	İ	ĺ				LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
						AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
1	ĺ	ĺ	1		!	EGFEWOLKLYQAMGYEVDTSSAIYKQYRLQ
		ļ				KVTEKYPELQNLPQELFAVDPTTVSQGLKDE
		1				VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
1 :	1	i		•		KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
]] '	ļ			LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
1						GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	A	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
] ·]	•	j]			TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
1			,			LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
1						GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
]		j				PTANREINPGPAAAADTRSCWGHKRSWRGW
1						RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
1						KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
						VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
			ŀ			HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA
						*AASTPKRAASIAHNAVSCR*AQIA
360	1710	Ā	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
					!	AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
]						*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
						QA
361	1711	Α	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
•						AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
						VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
						GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
				-		GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
						PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
						APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
						GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
						PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
					1	SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
			·			GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
	1					F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR
L		L	L			QG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RNQSPLGNDTLSSGLPMGPRRQVWPLARVG GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
362	1712	A	3136	1270	274	DNL RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMMMQTAGL EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA PAMNESPLAPHLHOHLVFSVFOVLTILIGV**
364	1714	A	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI SGASLCPVLSPPR*PLQPSSL
365	1715	A .	3145	122	413	LLPYPSLFVFLRQCHFVT/RLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSIPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPI.HSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVPMSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	Α	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*1*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVS\S*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	A	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410		RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSENPPI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l	ļ .		peptide		/=possible nucleotide deletion, \=possible
]		sequence		nucleotide insertion
		· -				FEIR\MYDAQHQQVGSNKCRVNNAGCSSLCL
†	ł	1				ATPGSRQCACAEDQVLDADGVTCLANPSYVP
	}					PPQCQPGEFACANSRCIQERWKCDGDNDCLD
İ						NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
1	ł	1				LCDGDNDCGNSEDESNATCSARTCPPNQFSC
						ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
]	!!!				CFPLTQFTCNNGRCININWRCDNDNDCGDNS
!						DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
	İ					NDCGDYSDETHANCTNQATRPPGGCHTDEF
	}					QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
		[GVTHVCDPSVKFGCKDSARCISKAWVCDGD
						NDCEDNSDEENCESLACRPPSHPCANNTSVC
						LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
						GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
1						CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
]				WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
	ļ '					KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
	1	·				VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
						LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
						TLLAGDIEHPRAIALDPROGILFWTDWDASLP
ļ						RIEAASMSGAGRRTVHRETGSGGWPNGLTV
1						DYLEKRILWIDARSDAIYSARYDGSGHMEVL
						RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
		1				KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
						ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR
						GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
						QRVYWSDVRTQAIKRAFINGTGVETVVSADL
					·	PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
						RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
						WTDGDNISMANMDGSNRTLLFSGQKGPVGL
						AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
1						AMRSQLGKATALAIMGDKLWWADQVSEKM
						GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
						QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
						MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
						GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
! '						TIYWVDMGLSTISRAKRDQTWREDVVTNGIG
						RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
			1		ŀ	SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
					ļ	GQYPRIERSRLDGTERVVLVNVSISWPNGISV
					, 1	DYQDGKLYWCDARTDKIERIDLETGENREVV
		l			-	LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
•			1			RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
		.				QKGTNVCAVANGGCQQLCLYRGRGQRACA
			ľ	ĺ	(CAHGMLAEDGASCREYAGYLLYSERTILKSI
		ŀ				HLSDERNLNAPVQPFEDPEHMKNVIALAFDY
		1	ŀ			RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT
		i i	ľ			IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
					ŀ	RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL
			J			TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
			1			RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
						WTDWVRRAVQRANKHVGSNMKLLRVDIPO
		١ ١	Į	ļ	ļ	QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
				[LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR
'		.	1			AQDEFECANGECINFSLTCDGVPHCKDKSDE
		·	ł	ł	ŀ	KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
		Į	1	ļ		GADDCGDGSDEIPCNKTACGVGEFRCRDGTC
		j	ļ	İ	İ	IGNSSRCNQFVDCEDASDEMNCSATDCSSYF
					 	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND
						CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
						MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
•						CONHRCISKOWLCDGSDDCGDGSDEAAHCE
						GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
						DGADESIAAGCLYNSTCDDREFMCQNRQCIP KHFVCDHDRDCADGSDESPECEYPTCGPSEF
						RCANGRCLSSRQWECDGENDCHDQSDEAPK
1						NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
						GODDCGDSSDERGCHINECLSRKLSGCSQDC
1						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
						TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
}						HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
1 1						TLLKQGLNNAVALDFDYREQMIYWTDVTTQ
						GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
						DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
						VSSGLREPRALVVDVQNGYLYWTDWGDHSL
i						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
					1	RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTTGTN
						KTLLISTLHRPMDLHVFHALROPDVPNHPCK
						VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE
						DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
						PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
1		ĺ				CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
					Ï	KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
		1				KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
						PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEVDCTPRCDMDOEOCKSCHCPL BWRCDA
						DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT
						LCKPLAWKCDGEDDCGDNSDENPEECARFV
1						CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
1 1		ľ		į		GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
1		- 1	į			SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
1 . 1						NASICGDEARCVRTEKAAYCACRSGFHTVPG
						QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
				ļ		ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
						LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA
		ĺ	[GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR ROIDRGVTHLNISGLKMPRGIAIDWVAGNVY
				l		WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
		1				AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
1						LRETLYODNIOWPTGLAVDYHNERLYWADA
1		J	}	J	j	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
	1				l	FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG
					ļ	LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
					1	CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
					ļ	PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
1	l	ŀ	ı		i	QPRYTGDKCELDQCWEHCRNGGTCAASPSG
		1	ļ			MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
		ł	1			VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
	Į					NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
]	1					CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECOCPPHM
1	[ĺ	[TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
	ļ		ļ			VAGVVFWYKRRVQGAKGFQHQRMTNGAM
1		1	ļ			NVEIGNPTYKMYEGGEPDDVGGLLDADFAL
1			1			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DPDKPTNFTNPVYATLYMGGHGSRHSLASTD EKRELLGRGPEDEIGDPLA
374	1724	A	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL NSQTTTADQLLEFFKQVGEVKFVRMAGDET QPTRFAFVEFADQNSVPRALAFNGVMFGDRP LKINHSNNAIVKPPEMTPQAAAKELEEVMKR VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA K*FFPPRVWKEQ*KKRRISRSHTRSKSRSSSK SHSRKRSQSKHRSRSHNRSRSQKDRRRSK SPHKKRSKSRERRKSRSHSRDKRKDTREKI KEKERVKEKDREKEREREKEREKERGKN KDRDKEREKDREKDKEKDREREREKEHEKD RDKEKEKEQDKEKEREKDRSKEIDEKRKDK KSRTPPRSYNASRRSRSSSRERRRRSRSSSRS PRTSKTIKRKSSRSPPRSRNKKDKKREKERD HISERRERERSTSMRKSSNDRDGKEKLEKNST S
375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP WF\FFYP*SPDLQIPSSFRNGLNDWYHHSQKC PDLDGAYVKK
376	1726	Α	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K
377	1727	A	3201	274	1285	KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA+PSPDAKYNCH+*QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY
378	1728	A	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTTKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS
379	1729	Α	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK *LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion XKQSNNKGTRASSYXEPDAXEQWKFPHKKL
380	1730	A	3207	187	507	QLPGXTHE GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP AXLLPGPGGGPGPVASLEARAQASSGVTPNG GGRTYPYPTFSSGE
381	1731	Α	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/ EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS KAWRASQMMTFFIFLLFPSFTGVLCTLAITI WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI KLQDMEKKANPSSLVLERREVEQQGFLHLGE HDGSLDLRSRRSVQEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHHQGHNS LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF LFSNACL
383	1733	A	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI
384	1734	A	3242	3	678	IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI
385	1735	A	3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MTCADPGEIANGHRTASDAGFPVGSHVQYRC LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF
386	1736	A	3250	5725	3984	SNPLYEAGDTREYEVSI GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	Α	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	Α	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQEELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	Ā	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLILGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

[050 th	000 70	CXC	Lore	D. 4:4:3	I positive in the	[]
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
nuci- eotide	peptide		in USSN	location		F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
	seq- uence	1	09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uciice			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	j		peptide	Sequence	/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
			 	sequence		RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
			1			KOOO
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
		l	3200	1	1 5.0	WICLSMVILTHSLKTFHRNWDWESEYTLFMS
					•	ALKVNKNNAKLWNNVGHALENEKNFERAL
]		})	KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
			-	_		IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
		l				WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
		ŀ				AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
		<u> </u>			į	GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
		ł				GLT
398	1748	Α	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS
						MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
		I				STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
						LKGLYGNRDLTARLELFTGRFKDWMVSMIV
		ł	1		1	DREYSVAVEAVRLLILILKNMEGVLMDVDCE
		[į	SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
						RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
						SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
ĺ		ľ				VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
						VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
		l	1			SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
						CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
					}	PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
						GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
•	1	Ī				STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ
						SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV
		j				LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
1		ŀ	1		}	SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
						N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
						TGAALAGSYPIWENENTLSWLPTFTYNFCLST
		ļ	Į j			PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
1		1	i			NILPPNQTILISVEASISSSPIRNKWALHLITLLT
						GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
		l				DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
						TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
						DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
		[LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
						ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
400	1750		2202		452	THURTIOGOUDOOTTA DIDUDEL ELA TENNOE
400	1750	Α	3303	2	453	THWRHSSGVPGSTTARRRRRELEIATSDNQE
	•					YYNRLCQEVTNRERNDQKMLADLDDLNRTK
						KYLEERLIELLRDKDALWQKSDALEFQQKLS
	•					AEERWLGDTEANHCLDCKREFSWMVRRHHC
40:	1751		2704		(2)	RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
						CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
						EKIVLRALKDSRAGMPEQDKDPGVQENPDD
						QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
						GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
						NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
402	1950	_	3205	1.000	170	LSLSTQILG*QKPSKYIPSLCKR
402	1752	Α	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
i l			1			PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
					1	LANDON ACALLA CILLAN CORROLO POLICIONO CONTROLO
						VVRTMCAVLGLVARQEDSGLRDHSVRVLISN HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ FFFFIFLNIFLLAFSSPGSQPLLNSPPSFVCWSR GFMEMNGRGELVESLKRFCASTRLPTPLLLF PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ VQRTLVSVTVSDASWVSELL\WSLFVPFTVY
403	1753	A	3307	44	447	QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE LG\QTGTRLTPA\DKAEHMKRQRHPR\LRPQS AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF CPHVAIGVFIPERPWPKTGCCKTLTIHLILL*G GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL QEKKQ\ALYEYARRFTERRAPGGLD DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS
403	1733		3307	11	447	GGASAGLASSPECACGRSHFTCAVSALGECT CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD CPPRECEED
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG QDHVQNEEIYARVLDKFGSNFLSRDNADLGT AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS LLKGDLKGVKGDLKKPFDKAWKDYETKFAK IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIK\IWE GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV MCVLLWALSLLQSILEWMFCSFLFSDVDSDN WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESFHLAKDSGFKVVAHMMP DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP TLVIRGTGLYELWKSGRYKSYSPSDLVELVA RILALVPPWTRVYRVQRDIPMPLVSSGVEHG NLRELALARMKDLGIQCRDVRTREVGIQEIH HKVRPYQVELVRRDYVANGGWETFLSYEDP DQDILIGLLRLRKCSEETFRFELGGGVSIVREL HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL QGPYMVKMLK
408	1758	A	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG LKGRVFEESLADLQND\TDGYLLRVI*VAFTT ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHOK
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLIP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLRDSLTSLGDYVLTCRWRNQALHFKIN KVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAHYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGVRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	RRYEKFDK VLTALSHKLEPAVRSSEL IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEPLTERIKYD PIPVRWNSLEGRLLRGYEQHANDGKDYISRN *DLRSWTAADMAAQITKRKWEAEEFAEQIKA
						YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	A	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	A	3398	304	2121	EEEEEEDEDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NKVHADLVISKPVSKSPERLRKDIEVLSEDTD YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR RYCNTEECLKTGSPGKKEEKAKNKESLCMEN SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE AAASPPHPAPEEGVAEESLQTVAEEESCSPSV ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
						KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK OOKEGK
419	1769	A	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP NGVVLDTQQDQLENAKMEHTNASFDTFFCE TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV VMGFHHVGQAGLELLTSGDLPALASQSARIT GVNHCAOPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKLHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPPTVILEPHVFP GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH
422	1772	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE
423	1773	Α	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL
424	1774	A .	3421	4	. · · · · · · · · · · · · · · · · · · ·	RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV LQISATDRDSGLNGRVFYTFQGGDDGDGDFI

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:			
		110G		beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	!	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ			peptide		/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
1						GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
1						
1						VFVEENSPIGLAVARVTATDPDEGTNAQIMY
						QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
1						YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
						LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
						DISDSLTYSFERGNELSLVLLNASTGELKLSR
1						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
						TITTDEMLTHSITLRLEDMSPERFLSPLLGLFIO
						AVAATLATPPDHVVVFNVQRDTDAPGGHILN
						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
						LLTAISAQRVLPFDDNICLREPCENYMRCVSV
1 1					†	LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1						TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
]]						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
1						YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
					,	TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
i i						QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
						VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
				ļ		VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
						ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
				İ		l
1 1						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
!				ļ		LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
						GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
i l						QASSLRLEPGRANDGDWHHAQLALGAIGGP
		1		1		GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
1 1						GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
1						SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
!						CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
						EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
1 1		1	ł			RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
i i				İ		I
	1					NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
						SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
1 1	l	I	1	l		AEVTTNGCEVNYDSCPRAŒAGIWWPRTRFG
] [Į		i	Ī		LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
] 1	I		İ	ļ		NCTSITFSELKGFAERLQRNESGLDSGRSQQL
1			j			ALLLRNATQHTAGYFGSDVKVAYQLATRLL
1		ŀ	ì			AHESTQRGFGLSATQDVHFTENLLRVGSALL
]]	J	ŀ	ļ	J		DTANKRHWELIQQTEGGTAWLLQHYEAYAS
] [[ŀ	İ	ĺ	ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
	ļ	ļ	[FAGAKLPRYEALRGEQPPDLETTVILPESVFR
	}		l	ļ	l	ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
j j)		}	J	j	GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
1 1	ļ	1	J	į		
1	ļ	ł	1	1	l	RPIINTPVVSISVHDDEELLPRALDKPVTVQFR
1 1	İ		1	1	l	LLETEERTKPICVFWNHSILVSGTGGWSARGC
į į]		ļ	j	J	EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
; I	İ	I	j	į		NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
ļ I	1	I	ļ	l		RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
1	}	l	l	ļ		DLPFACTVIAILLHFLYLCTFSWALLEALHLY
Į !	İ	j	j	į		RALTEVRDVNTGPMRFYYMLGWGVPAFITG
	ŀ		l	.		LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
į l		į	1	ł		VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1 1	ŀ		1	l	1	PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
1	ŀ	1	ł	ł	[-
1 1		j		•		FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
. 		1		ł	1	LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
[į		RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
j	}			ŀ		EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ
				1		H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EEEEEEEEAAFPGEQGWDSLLGPGAERLPLHS TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP EERLRENGDALSREGSLGPLPGSSAQPHKGIL KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE GSRGGPPSRPPPROSLQEQLNGVMPIAMSIKA
425	1775	A	3429	155	1417	GTVDEDSSGSEFLFFNFLH GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS RAASVREAEDAPLQPASIHPVSQGSRGPEGSL GSAECLPGDPLGARRATRAHSPVPGPPPSLPA AGTAVKRGLQPG*GA/GATSTPGTGAATGGL CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGMVILHRISAADPARAPGPDS NLQSALQQPATGCSEPAAVYSPPIGLWGA**P EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS
426	1776	A	3431	1662	369	AİWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTGSHDFAINFDPENPECEGK RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVFNEISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP
427	1777	A	3446 .	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPPQPPPPPPPPPP

SEQ ID SEQ ID Met SEQ Predicted NO: of NO: of NO: of nucl- eotide seq- Note of NO: of NO: of nucleotide location corresponding le-Isoleucine, K=Lysine, L=Leucine,	steine,
nucl- peptide in nucleotide location F=Phenylalanine, G=Glycine, H=Histid	
entide seg- USSN location corresponding lateraleucine Kallysine I al eucine	in e ,
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, P=Proli	ne.
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=Serine,	
amino acid of peptide T=Threonine, V=Valine, W=Tryptopha	n .
residue of sequence Y=Tyrosine, X=Unknown, *=Stop code	
peptide /=possible nucleotide deletion, \=possible	le
sequence nucleotide insertion	
TTEYPEEQYVSDILNYIDHGDPQVR	GATAILC
GTLICSILSRSRFHVGDWMGTIRTL	
ADCIPLLRKTLKDESSVTCKLACTA	
SLCSSSYSELGLQLIIDVLTLRNSSY	WLVRTEL
LETLAEIDFRLVSFLEAKAENLHRG	AHHYTGL
LKLQERVLNNVVIHLLGDEDPRVRI	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	_
IRLVPKLFYKCDQGQADPVVAVAR	•
KLLMHETQPPSHFSVSTITRIYRGYY	
VTMENNLSRVIAAVSHELITSTTRA	LTFGCCE
ALCLLSTAFPVCIWSLGWHCGVPPI	SASDESR
KSCTVGMATMILTLLSSAWFPLDLS	
	•
ILAGNILAASAPKSLRSSWASEEEA	
QEEVWPALGDRALVPMVEQLFSHI	
AHVLDDVAPGPAIKAALPSLTNPPS	LSPIRRK
GKEKEPGEQASVPLSPKKGSEASAA	SRQSDTS
GPVTTSKSSSLGSFYHLPSYLKLHD	-
NYKVTLDLONSTEKFGGFLRSALD'	The state of the s
	•
ATLQDIGKCVEEILGYLKSCFSREPA	
VQQLLKTLFGTNLASQFDGLSSNPS	KSQGRA
	TQALADA
SLRNMVQAEQENDTSGWFDVLQK	VSTOLKT
NLTSVTKNRADKNAIHNHIRLFEPL	-
, , , , , , , , , , , , , , , , , , ,	•
YTTTTCVQLQKQVLDLLAQLVQLR	
DSDQVFIGFVLKQFEYIEVGQFRESI	
	IMASGR
KAVTHAIPALQPIVHDLFVLRGTNK	ADAGKE
LETQKEVVVSMLLRLIQYHQVLEM	FILVLOO
CHKENEDKWKRLSRQIADIILPMLA	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
DSHEALGVLNTLFEILAPSSLRPVDN	
	•
	DSTSTLE
	LLEDIVT
KOLKVEMSEQOHTFYCOELGTLLM	
GMFRRITAAATRLFRSDGCGGSFYT	
1 1 1 1 1 1 1 1 1	
ARSMITTHPALVLLWCQILLLVNHT	
AEVQQTPKRHSLSSTKLLSPQMSGE	1
AKLGMCNREIVRRGALILFCDYVCC	(NLHDSE
HLTWLIVNHIQDLISLSHEPPVQDFI	SAVHRNS
AASGLFIQAIQSRCENLSTPTMLKKT	
HLSOSGAVLTLYVDRLLCTPFRVLA	
ACRRVEMLLAANLQSSMAQLPMEE	
LQSSGLAQRHQRLYSLLDRFRLSTM	` '
PPVSSHPLDGDGHVSLETVSPDKDW	
SQCWTRSDSALLEGAELVNRIPAED	MNAFM
MNSEFNLSLLAPCLSLGMSEISGGO	
AREVTLARVSGTVQQLPAVHHVFQ	
AAYWSKLNDLFGDAALYQSLPTLA	
LVVVSKLPSHLHLPPEKEKDIVKFV	VATLEAL
SWHLIHEQIPLSLDLQAGLDCCCLA	LOLPGL
WSVVSSTEFVTHACSLIYCVHFILEA	
EOLLSPERRTNTPKAISEEEEEVDPN	
1 1 1 1 1 1 1 1	` ,
TAACEMVAEMVESLQSVLALGHKF	
	LVWKLG
WSPKPGGDFGTAFPEIPVEFLQEKE	/FKEFIYR
INTLGWTSRTQFEETWATLLGVLVT	
QEESPPEEDTERTQINVLAVQAITSL	
VPVAGNPAVSCLEQQPRNKPLKALI	
LSIIRGIVEQEIQAMVSKRENIATHHI	
PVPSLSPATTGÅLISHEKLLLQINPER	ELGSMS
YKLGQVSIHSVWLGNSITPLREEEW	

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of, peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
						LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	Α	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA SGETDSE
430	1780	À	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEVHER\SVQS DFLLIILKEILQKRSDLHLILMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
431	1781	A	3474	I	-	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEQ ID NO: of nucl- eotide seq- ucnce	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Argininc, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
						QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS APRSRCVARPAARTGLPTPAPASSPAPAASPA PAASPAPAASTA\PQPLILLPKP/PPAPGAPPPRP GAPPPRPAASPSPAASPAPLTASPPLP AASPSPAASPAPPAASPVLTASPPLPAASPSPA ASPAPPAASPVLTASPPLPAASPALAASPVHT ASPPVHVASPPVHTASPPLPAASPVHTASPP VHVASPPVHTASPPVHVASPPVHVASPPVHV ASPPVHTASPPVHVASPPVHVASPPVHVASPPVHV ASPPVHVASPPVHVASPPVHVASPPVHVASPPVHV ASPPVHVASPPVHVASPPVHVASPPVHVASPPVHV ASPPVHVASPPVHVASPPVHVASPPVHVASPPVHV ASPPVHVASPPVHVASPPVHVASPPVHVASPPVHV
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG
	1785	A	3529		3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFA VEDSQLVELTKQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQDDILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAAATY VQEFLPHSICAMQ
436	1786	A	3546	73	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion KGKGKTIRGI*TFKGRKGGTYQREHDANPLA PXSARSCWMRKG
437	1787	A	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG SASGCRYAAHPHWGLGGAAAAGGSWEPQPP RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA AQPGARARTSPPPASARNMAARPAATLAWSL LLLSSALLREGCRARFVAERDSEDDGEEPVVF PESPLQSPTVLVAVLARNAAHTLPHFLGCLER LDYPKSRMAIWAATDHNVDNTTEIFREWLK NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDQAQLDWELIYIGRKRMQVKEPEKA VPNVANLVEADYSYWTLGYVISLEGAQKLV GANPFGKMLPVDEFLPVMYNKHPVAEYKEY YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLSTLTGRSY GQPSLQDELKDNTTVFTRILDRLLDGYDNRL RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI DVFFRQSWKDERLKFKGPMTVLRLNNLMAS KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH ACPLKFGSYAYTRAEVVYEWTREPARSVVV AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK FWEVISDEHGIDPTGTYHGDSDLQLERINVYY NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569		1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ FIEELITKWQKNDQELISDPLQQCFKKDEILDG QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG QDRLAVLLPGRHPCDCLGQKHKLINNCLICG RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS NKSQKLLKKLMSGVENSGKVDISTKDLLPH QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid	to last amino acid residue of peptide	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						DDESDYFASDSNQWLSKLERETLQKREEELR ELRHASRLSKKVTIDFAGRKILEEENSLAEYH SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
						VNPNMYQSPPQWVDHTGAASQKKAFRSSGF GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
						TAKKPSPQEVSELQATYRLLRGKDVEFPNDY PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
		<u> </u>				QGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMRLSSLTRHMRSHT AIRANEKPYKCKEC\GRAFSLSQILSK\HERSH
,		,	:			TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK PYECKQCGKALSCSSSLRVHERIHTGEKPYEC
		ļ				KQCGKAFSCSSSIRVHERTHTGEKPYACK\EC GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
						FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
						TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
						HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG EKPYECKECAKTFISLENFRRHMITHTGDGPY
						KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
		'				CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
					•	LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
†						QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
						PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
						DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
443	1793	A	3578	287	114	ANTVKPHLY DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK
						LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV
	ĺ					MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
						RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA
		İ				DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL
]				GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
						KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
						KLCDWWSLGVIMYEMLIGYPPFCSETPQETY
						KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
ļ						AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
	}					KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK
						NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
						FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE
						KTMENLKQGQFTHTRNQPKGVTQIINYTTVA
						LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA
	1					FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP TNVEDVCPNIPSLEKLMEEIVELAESGIRYTO
L	<u> </u>	<u> </u>	L		l	114 TEN VCI MIT SCENE MEET VELACOURT IQ

seq- uence 914 914 916 10 to last amin acid residue of peptide peptide sequence 917 918 919 919 919 910 910 910 910	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence 9149 gr to first	nucl-	peptide			nucleotide	1	F=Phenylalanine, G=Glycine, H=Histidine,
Service Serv							1
amina add residue of peptide requence peptide sequence peptide sequence peptide sequence peptide sequence per		uence			-		, , , , , , ,
residue of peptide sequence	Lonoc			717			
sequence nucleotide insertion							
MPHYMEVILPMIC SYMSRW WEHGPENNER ARMCTAL NISEMBNITLIGHILLIUNING GIB GAWMRLAYFSQPIINKVRPOLLKTHFLPLM KELKKKAATVVSEEDHIKAEARGMSEAEL LILDEFTILARDI. YAPYPLLIRFGDYNRAKWLL KEPNERAELERMYAEVFYWSKSINHSKREE QNFVVQNEINIMSFLITDTKSKMSKAAVSDO ERKKMKRRGRPSYMOTSI. IVAALKELIFIGI NICAPGDQELIAL AKNRTSI. KDTEDEVEDIURS NIHLQGKLEDPARWMALYKDLPHSTDDTS DPEKTVERVLDANVL FHLEDKSKRVQRIIIY CLVEHPOSKKAVWKLISKQRKRAVVACF RMAPLYNLPRHRAVNLELOGYTKSWHITEEH YEDKLLEDLAKVGAEPEEDBCSTERVPLHQ LILLISRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDCIEEVKSFEKEMEKOKLLIQQ ARLHDRGAAEMVLQTISMSKGETGPMVAAT LKLGIALINGGNSTVQKMLDYLKKKOVLY FQLAGALMOSCSYLDLAMFERONKASILLOH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIFRLQLLCEGH WTEEGSGEKVLQDDEFTCHIRGLLCGLC KDMVYMLLSTLEGHOWYDLLV ESISPFWYTYSGKOMEGOGGRAFEKLDVAV VGFLHYFAHMQMLSQDSSQIELKELDDEFTKETYD PDGKGVKRADFIKAMESIKHTYGSTETELL SCAETDEWETLDYEEVKRFHEJPAKDIGFNVY VLINISSHMPONTRUTHLELAESVLNYQOP FLORIEMGSAKRIBRVYFEISSSRRQWEKAP VKESKAGHFDVVNEGGREKEMELFVNFCDE TIFEMQLAAQISSINLHERSINKSESKERPEE QOPPMARTSILTYRSALPARVINLITHMEMLS LKSLKQMKKVKMTVKDMVTAFFSSYWS FMTLLHFVASVFROFFRICSLLIGGSLVEGA KKIKVAELLAAMPDPTQDEVKRODEGGERER LEAALFSSLTINKALFARNYNILTHMALLS LKSLKQMKKVKMTVKDMVTAFFSSYWS FMTLLHFVASVFROFFRICSLLIGGSLVEGA KKIKVAELLAAMPDPTQDEVKRODEGGERER EGGGYKLIPHNPAGLISDLMSNIPVRWEVQE KFGRQKAKERGEREEKERFEKSEFKAEGGEGGE KEEKAKEDSGKKKLGLITHTRSSETAH RYKLDSSSRRIAAVFVLEESSCYMPETVRLL PHLHTVISFFCIIGYCLKVPLUFFREKEVARK LEFDGLYTHOYSFFCINGTON TOSPF NNYWDKFVKRKWMDKYGFFYORDNISELLG MDKAALDFSAARVANGFFTYHTUSSYTH NGKQLVLTVGLLLAVVYLVTVVAFNFFRKF YNKSDOLTDMKCDOLMITCYMFHWYGY RAGGGIGDEEDPAGDEVETYRIIIFDTTFFFV VILLAIDGLIIDAG					- •		, · · , · ,
ARMCCTALNSEHMNTLLGNILKIJYNNLGIDE GAWMRALAYSOPIUNKVPQLLKTHFIPLM EKLKKAATVVSEEDHLKARAGOMSEAEL LILDEFTTLARDI,VAPPULIRFODYNRAKWL KEPNPEAEELFRMYAEVFTYWSKSHNFKREE ONFVOYDEINMASHLTDTYSKMSKAAAVSDO EKKKMKRKGRYSMOTSLIVAALKRLLPIGL NCAPODOELIALAKNRTSIKDTEDEVBUIRS NHLQCKLEDPARWOMALYKOLPNTDDTS DPEKTVERVLDIANVLPHLEQKSKRVGRRIYV CLVEHPORSKKAVWHKLISKORKRAVVACE RMAPLYNLRRIRAVNLPLQGYEKSWIETEEH YFEDKLEDLAKPGAEPFEDEGTKRAVDPHQ LILLFSRTALTEKCKLEEDFLYMAAT LKIGHALINGINSTYQGKMLDYLKEKKDVGF FOSLAGLMOSCSVLDLNAFERONKAEGLGM VTEGSGEKVLQDDEFTCDLFRFLQLLGEGH NSDFQNYLRTGGRNTTVNIHISTVDYLLRVQ ESISDFYWYSGKDVDEQGQRNFSKALOVA KQVFNHTATYQGPCTGNQOSLAHSRLWDAV VOFLHVPAHMQMKLSQOSSQIELKEMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMMLKFFDMFLKLKDLTSSDTYFEYD POGKOVERKPFHKAMESHKHYTQSETEFLL SCAETDENSTLDVSEFEYKRFHERAKDGRNVA VULTNISEHMPNDTRLQTFLELAESVLNYFOP FLGREIMGSAKRIERVYPEISESSRTOWEKPQ VKESKRQFIFDVVNEGGEKEKEMELFVNFCED TIFFMGLAAQISESDLAKSKREPEE QOPPMAFFSILTVRSALFALRYNILTHMMALS LKSLKKOMKVKKKMYKNDWTAFFSSYWSI FMTLLHFVASSVFROFFRICSLLIGGSLVEGA KKIKVAGHLANDPOTPODEVEGGEGEREK LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYKLPHPNFAGLSDLMSPROPEVPESA FWKKILAYQOKLLNYRANFFNNMMLALFUR LKSLKKOMKKVKKMYKNDWTAFFSSYWSI FMTLLHFVASSVFROFFRICSLLIGGSLVEGA KKIKVAGLANMPOTPODEVEGGEGEGERE LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYVLPHPNFAGLSDLMSPROPEVPESA FWKKILAYQOKLLNYRANFFNNMMLALFUR LEFDGLYHTPOPSEDDIKGOWORLVINTOSEP NNYWKKYKWKWNDWTAFFSSYWSI FMTLLHFVASSVFROFFRICSLLIGGSLVEGA KKIKVAGLANMPOTPODEVEGGEGEGERE LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYKLLHYNFOFTODEVEGGEGEGERE LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGGYCKLPHTHRYGGTEVPESS PNYKKILAYQOKLLNYRANFFYNMMMLALFV AFANFILLYFYKTSSVVOCKEL PTRSSSENA KYTSLDSSSHRIAVHYVLEESSGYMEPTVRIL PLHTVINSFFCIGHYVCLKVYVINFREKEVARK LEFDGLYTHPOPSEDDIKGOWORLVINTOSEP NNYWKKFYKEKWMOKYOFTORSPROF NNYWKKFYKEKWMOKYOFTORSPROF NNYWKKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWMOKYOFTORSPROF NNYWKFYKEKWOMOKYOFTORSPROF NNYWKFYKERO					sequence		
GAWMKRLAVFSQPIINKVRPQLLKTHFLPLE EKLKKAATVVSEEDHILAGEAGMSEAEL LILDEFTTLARDI.VAFYPLILREGDYNRAKUL KEPNPEAEELFRWAVEVFIWSKSINFKREE QNFVVQNEINNMSFLITDTKSKMSKAAVSD EKKKAKRKGRDR SAMGYSLIVAALKRLIPIGL NICAPGDQELIALAKNRFSLKDTEDEVRDIRS NIHLQGKLEDPAIRWQMALYKDNFTDDTS DPEKTVERVLDIANVLFHLEQKSKRVGREIVI CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLVEHPGRSKKAAWHKLISKORKRAVVACY CLUEHPORSKCAWHKLISKORKRAVVACY CLUEHPORSKCAWHKLISKORKRAVVACY CHENDEDDGGEEVKSFEREMER KOLLTYN LILGIALINGGRINDOW CONTROLLED THESCHENDER CONTROLLED CHENDER VYESSGEVULDEDGGGREKARELMDLA VOGHAFFSILTVARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLA VILLTILSEHMPRDIFUKARFREPARDIGFRLAMMILSEVANTOR VILLTILSEHMPRAGLACHAWHTAFTSSTVSN FINGAAQUSESDLAFERSANKESSEKERFRE QOPRAFFSILTVARSALFALEVNITARSSTVA VILLTILSEHMPRAGLACHAWHTAFTSSTVSN FINGAAGAGGEGEERMETKSSESSERTOWERS ERSCANGENGARGERKENTSSESSERTOWERS VILLTILSEHMPRAGLACHAWHTAFTSSTVSN FINGACHARPSTRATTAR LEAALTSCORT TARBOTT LEAGUS TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW TO THE MORROW	1	'					
EKI, KKKAATVVSEEDHLAKARGOMSEAL LÜLDEFTTIARDI/AFYPLIRGOPVINRAKWL KEPNPEAELIFBMVAEVFIYWSKSHINKRISE QNEVVONDININMSHITDITSKSIKSKAAVSDQ EKKKMKRKGDRYSMQTISLIVAALKRIL PIGI NICAPGOQELIALAKNEPSILKDTEDEVRDIIRS NIHLQGKLEPPARWQMALYKDLPNRTDDTS DPEKTYVER VUDIANVIHLEQKSKYGRIIVY CLVEHPQRSKKAVWHKLLSKORKRAVVACF RAMAPLYTHPHRAVNILLGOYEKSWTETEH YEROKLEDLAKPGAEPPEDEGTRAVDPHAQ LILLFSRTALTEKCKLEEDFLYMAYADDMAKS CHDEDDDGEEEVKSPEREMBERGKLLVQQ ARLHDRGAAGMVUQTISASKGETGMVAATI LILLGIALINGGNSTVOQKMLDYLKEKKDVGF FOSLAGLMQSCSVLDLNAFERQNKAEGLIOM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDPQNYLRTQTGNNTTVNIISTVDYLLRVQ ESISDFYWYYSKGWDDGQORPSKAIQVA KQVFNTLTEVIQGPCTGNQSLAHSRLWDAV VGFLHVAHMGMKLSQSSSQELELMDLQ KDMVVMLLSMLEGRVVNGTIGKQMVDMLV SSNNVFMLKFPMFMFLKLDLISSDTFKEYD PPGKGVFKRDFHKAMESHKHYTQSETEFLL SCAETGENETLDYBFMFLKLDLISSDTFKEYD PPGKGVFKRDFHKAMESHKHYTQSETEFLL SCAETGENETLDYBFFVKRTPHFAKDGFNVA VLLTNLSEHMPNDTRLQTFLELASSVLNYPQP FLGRIEMGSAKREVYFEISESRTWBKPQ VKESKROFIFDVVNEGGEKKMLEVNTCED TIFFMQLAQDISSDLNERSANKESEKERPEE QGPRNAFFSILTVRSALFALRVNLTLHMBLLS LKSLKKKMKKVKKMTVMDMTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGG KIKKVABLLANNPPTQDEVRGDEGERKP LEAALPSEDLTDLKELTESDLSDIFGLDLK GGGVKLIPHNPNAGLGDLMSNPVMPPEVQE KFQGCKAFEEKEEERSFFKAEGGEGE KEKAKEDKGKGKLRQLHTHRYPEVPEVPLOF FMFLLHFVASVFRGFFRICSLLLGGSLVEGG KKKVABLLANNPPTQDEVREPOYPESEA FWKKILAYQQKLLNYFARPYNMRMLALFV AFANNFILLFYKVSTSSVEGKELFYRSPENDEVPPEND FWYSSKRIPPTOPPENDEGGERKP LEAALPSEDLTDLKELTESDLSDIFFGLLKL CEGGVKLIPHNPNAGLGDLMSNPVMPPEVQE KFQGCKAFEEKEEKEFKSPFKAEGGEGE KEKAKEDKGKGKLRQLHTHRYPEVPEYPES FWKILAYQQKLLNYFARPYNMRMLALFV AFANNFILLFYKVSTSSVEGKELFYNRESENAA KVTSLDSSSHRILAVHYVLEESGGVMEDTVINSVL GHYNNFFFAAHLLDLAMGFKTLTRSSVH NGKQULTVTUNSFFLUGGERVP LEAALPSDEDDIGGQWDRLNINTOSSENA KVTSLDSSSHRILAVHYVLEESGRYREDEVPLA LEFDGLYTTEOPSEDDIGGQWDRLNINTOSSENA KVTSLDSSSHRILAVHYVLEENGERVPRAC LEFDGLYTTTOPSEDDIGGGWDRLNINTOSSENA KVTSLDSSSHRILAVHYVLEENGERVPRAC LEFDGLYTTTOPSEDDIGGGWDRLNINTOSSENA KVTSLDSSSHRILAVHYVLEENGERVPRAC GGGGGDBEDPAGDESURFINEIDTTFFFFY NNYWGKFVERTWANGERUNDENGERNEELLG NDKAALDTSPFTYNSVUCKERLTRTSLSVH NGKQULTYTTVSLFFTANGVLC FFTAGGGGGGG							Į.
LILDEFTILARDI.VAFYPILIRGDVINAKWI KEPNPEAELIFRWAEVFIYWSKSINFKREE QNEVVQNEINNMSFILITDTKSKMSKAAVSID ERKKMARKGRUS XMQTSLIVAALKRILPIGE NICAPGDOELIALAKNRESIKDTEDEVDIBLIS NIHQGKLEDPAIRWQMALYKDUPNETDDTS DPEKTVERVLDIANVLFHLEQKSKRVGRRIVY CLVEHPQRSKKAVHKLISKQRKRAVVACC RMAPLYNLPRIMTAVNLFHQUSKSWIETEH YFEOKLIEDLAKPGAEPPEEDGTKRVDPHHO LILLFSRTALTEKKLEEDFLYMYADIMAKS CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETWAVADIMAKS CHDEEDDDGEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETWAVATI LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQQKMLDYLKEKRDVGF FOSLAGLMGSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQCKMLDYLKEKRDVGF FOSLAGLMGSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQKMLDYLKEKRDVGF ESISDFYWYYSGKDVIDEGQGRNSSKAIQW KQVFNTLTEVIGGCGTCTONQGSLAHSLWDAV VGFLHVRAHMGMKLSQDSSQIELLKELMDLQ KDWYWLLSMLEGGTVNOKYDMLUN ESSNIVEMILKFFDMFLKLKDLTSSDTIFLEDL SCAETDENETLDYEEFVKRPHEPAKDIGFNVA VLITILSEHMPNDTRLQTHELGESVLNYFOP FORGVENFAREPHAKDIGFNVA VLITILSEHMPNDTRLQTHLGTENVYNTOP FORGVENFAREPHAKDIGFNVA VLITILSEHMPNDTRLQTHLGTHYOPPCED TIFFMQLAAQISESDLNERSANKESESKERPEE QOFRNAFFSILTVRSALFALEYNITLINKMUL SKLKKGMKKWKKMTVKMVTAPSSTVNS FMTILHFVASVFROFFRICSLLLGGSLVEGG KKKVABLLANNPDFTQDEVRGDEEGERKP LEAALPSEDLTDLKELTEESDLSDIFGOLDK GKKKVABLLANNPDFTQDEVRGDEEGERKP LEAALPSEDLTDLKELTEESDLSDIFGOLDK GKKKVABLLANNPDFTQDEVRGDEEGERKP LEAALPSEDLTDLKELTEESDLSDIFGOLDK GKKKVABLANNPTVMPEVOR KYGGCKALFEEKEETSSPEKAEGEDGE KEKKARKDKOKGKLRQLHTHRYSTSENDA KVTSLDSSSHRIAVHYVLEESSGYMEPTYVIL PHINTISFFCIGVYCLKVPLVIKFREKEVAR LEFDGLVTIEGPSEDDIKGQMDRLVINTOSSENA KVTSLDSSSHRIAVHYVYLEESSGYMEPTYFF NAWGKGKKLFRUTHSLLSTVH NGKGLTYTOLTHSLLTLANNL GHYNNPFFAAHLLDLAMGFKTLRTILSSTVF NAWGKGROFFTONSLVLAAVATSRLTAALFS*NQAWCL GLEKERPATAT]						, , ,
ONFVVONEINMSFLITDTISKMSKASKAYSDE ERKKMERKGRUP SMOTSLIVAAL IRLL PIGE NICAPGDOELIALAKNIPSIKDTEDEVRDIESS NICAPGDOELIALAKNIPSIKDTEDEVRDIESS NIHLQGKLEDPIARWOMALYKDLOPRITDDTS DPEKTVER VI, DIANVLPHLEQKSKEVGERIY CLVEHPQRSKKAVWHKLLSKQRKEAVVACY RAMPLYNLPHIRAVNLFLQOTEKSWIETEEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLH LILLESRIATERKEKLEEDLYMAY ADIMAKS CHDEEDDDGEEVKSPEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETWAVAT LKLGJAILNGGNSTVQUKMLDYLKEKKDOG FOSLAGLMQSCSVLDLNAFERGNEWAVAT LKLGJAILNGGNSTVQUKMLDYLKEKKDOG VTEGGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFONVLRTQTGNOTNTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEVIGGOFTCTNQGSLAHSRLWDAV VGFLHVFAHMGMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGHVVNGTIGKGWKMDVMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEVD PGKGVFKRDFHKAMESHKHYTQSFTEFLL SCAETDENETLDYEEFVKRFHEPAKDIGNVVA VLITALSEHMPIDTRLQTHELTEAVLNYPQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VESKRQFIFDVVHGGGKEKKEHFVPCED TIFFMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLARKLLS LKSLKQMKKVKMTVEMMTAFFSSTVNS FMTLLHVASVFRGFRRICSLLLGGSLVBGA LEAALPSEDLTDLEELTESDLSDIFGLDLKE EGGQVKLJHNFNRANGLSDLMSPYPMEVQE KKRVABLLANMPDTTOGDEVRGDEGEGERKP LEAALPSEDLTDLELTEESDLSDIFGLDLKR EGGGVKLJHNFNRANGLSDLMSPYPMEVQE KPQEGKAKEEKEEKESTKSPFEKAEGGDEG KEKKAKDKGKQKLRQLHTHKYGFEPEVPESA FWKKILAYQQKLLNYFARNYNMRMLALFV AFANNFILLPYKVSTSVYGKGERTMSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIL PHLHTVISFFCIIGVYCLLVPLIVERSESENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIL PHLHTVISFFCIIGVYCLLVPLIVERSEENAR LEFDGLVITIGPSEDDIKGQWDRLVINTOSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIL PHLHTVISFFCIIGVYCLLVPLIVERSEENAR LEFDGLVITIGPSEDDIKGQWDRLVINTOSSENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIVI AFANNFILLPYKVSTSVYGKGERTISESENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIVI AFANNFILLPYKVSTSVYGKGERTISESENA KVTSLDSSSHRILAVHYVLEESSGYMEPTYVIIVI VILLAINGDETBIFTFFTVIIVI VILLAINGDETBIFTGOFSTVCMWATAVIITISESTFT NGKQLVILTVGILLAVAYINTAVAVATANTAVIITISESTFT NGKQLVILTVGILLAVAYINTAVAVATANTAVIITISESTFT NGKQLVILTV							
ERKKMERGEN YSMOTSLIWAALKRILIPIGI MICAPGODELIALAKNERISLEDEVRDINES NIHLOGKLEDPAIRWOMALVKOLIPVETIDTS DPEKTYERVLOLANVEHLEGKYNEVGERIY CLVEHPORSKKAVWHKLLSRORKRAVVACF RMAPLYNLPRIRAVALELOGYEKSWIETEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDHIAKS CHDEEDDDOEEEVKSFEREKEMEKOKLLYOQ ARLHDRGAAEMVLOTISASKGETGFMVAAT LKLGIALINGGNSTVOQKMLDVLKEKKDVGF FOSLAGLIMOSCSVLDLNAFERONKAEGLIGM VTEEGSGEKVLODDEFTCHFOLLCEGH NSDFONYLRTOTGNNTTVNIIISTVDLYLRVO ESISDFVWYYSGKOVIDEGGGISHCPLLCLEGH NSDFONYLRTOTGNNTTVNIIISTVDLYLRVO ESISDFVWYYSGKOVIDEGGGISHCALEGLIGM VGPENTLTEYIQOPCTGNOQSLAHSRLWDAV VGPLHVFAHMOMLSODSSOJELLKELMDLO VGPHVFAHMOMLSODSSOJELLKELMDLO VKSKROPIEDVVNGGGEGERKMELFVHFCED TIFFMQLAAQISESDLURGSNAVERSVWS FMTLLHFVASVFGFRICISLLGGSLVFGA KKKVAELLANMPDPTODEVRGDGEGERKP LEAALPSEDLTDILKELTEESDISSTOJENGLIKA LKSLKKOMKKVKKMTVKDMVTAFPSSVWSI FMTLHFVASVFGFRICISLILGGSLVFGA KKKVAELLANMPDPTODEVRGDGEGERKP LEAALPSEDLTDILKELTEESDISFOGLDLKR EGGOYKLIPHNPNAGLSDLMSNPVPMPPVOPE VGPHVAHMOMLSODSSONAPSVWSINGSPHONALDFV AFANIFILEFVKVSTSSVVGKKEPTRSSSENA KVTSLDSSSHRIAVHYVLESSGYMEDFIVRI PHLIHVISFCIGIGTYCLKYPLIVFCED NVOMPKLGVVCTNONSTLYNGSPP NNYMOKFVKRKKVMDKYGEFRAGEGEGGG KEGAKEDKGRKRKPKKDSSLSAVLNSID VKVOMPKLGVVTDNSFTLYVAFNFRKF VNKSEGGOTPDMKCDDMLTCVMFHMYVGV RAGGIGGEBEPPAGDEVEIFFTFFFVI VILLAIGGLILNNSDPFALASGSGGTTGVTRTPSLFFF DTVLLCKGGBFALANDLINGFRINALPSVAORVCL SLPRSWDYRRWPHPHANFCTERDESL							KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
NICAPGOQELIALAKNIPSILKOTIBES WIHLQGKLEDPAIRWQMALYKDLPNIRTDOTS DPEKTVERVLDIANVLFHLEQKSKRVGRIFY CLVEHPORSKAVWHKLLSKORKRAVVACF RMAPLYNIPRHRAYMLELQCYSERSWIETEEH YFEDKLIEDLARGAEPPEEDECTKRVDEHQ LILLISSTALTEKCKLEEDPLYMAYADIMAKS CHDEEDDDOEEEVKSFEEKEMEKQKLLVQQ ARLHDBGAAEMUQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDVLKKKNOFG FQSLAGLMOSCSVULDNAPERMEKQKLLVQQ ARLHDBGAAEMUQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDVLKKKNOFG FQSLAGLMOSCSVULDNAPERMEKQKLLVQQ ARLHDBGAAEMUQTISASKGETGFMVAAT LKLGIAILNGGNSTVQOKMLDVLKKKNOFG FQSLAGLMOSCSVULDNAPERMEKQKLLVQQ ARLHDBGAAEMUQTISASKGETGFMVAAT CKLGIAILNGGNSTVQOKMLDVLKKKNOFG FQSLAGLMOSCSVULDNAPERMEKQKLVQQ VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQMYLTGGNTTVDVLLRVQ ESISDFYWYSGKDVUDEQGQRNESKAIQVA KQVFNTLTEYIQQFCTGNQQSLAHRBLWDAV VGFHVFAHMQMKLSQDSSQIELKELMDLQ KDMVVMLLSNLEONVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLLKLKDLTSDUTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFEDPAKDIGFNVA VLLTNLSEHMPNDTRLQTIFLELAESVLWTYQP FLOGRIEMGSAKRIENYYFEISESSRTQWEKPQ VKESKRQFFDVVNEGGEKEKMELEVNFCED TIFFRQLAAQJSSSDLNERSANTQWEKPC VKESKRQFFDVVNEGGEKEKMELEVNFCED TIFFRQLAAQJSSSDLNERSANTQWEKPCED TIFFRQLAAQJSSSDLNERSANTQWSKPQ VKESKRQFFDVVNEGGEKEKMELEVNFCED TIFFRQLAAQJSSSDLNERSANTQWSKPQ VKESKRQFFDVNEGGEKEKMELEVNFCED TIFFRQLAAQJSSSDLNERSANTQWSKPQ VKESKRQFFDVNEGGEKEKMELFVNFCED TIFFRQLAAQJSSDLNERSEKERPEE QGPRMAFFSILTVRSALFALRYNUTLARMLS LKSLKKQMKVKAMTVKDMTAFFSSVWSI FMTLLFFVASVFRGFFRICSLLLIGGSLVSGA KKIKVAELLANMPDPTQDEVRGDEEGERFP LEAALPSEDLTDLKELTEESDLSDIGTGLDLKR EGGQYKLIPHNPNAGLSDLMSNPPVMBFVQE KFQCKAKEERKEEKEETKSPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGGPFVESA FWKKILMYQGKLINYFANFTYMMRMLALFY AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAVHYVLESSGYMEPTVRL LFDGLYTTEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLO MDCAALDFSDAREKKRKNDSLSAVLNSID VKYQMWKLIGVYTDNSFLYLAWYMTMSVL GHYNNFFFAAHLDLAMGFFTLRTLISSVYH NGKGLULTVGLLAVVVLLYTVVAFNFFRFF YNKSEDGOTPDMKCDDMLTCYMFHMYNGV RAGGGGEDEDPAGDEVSTRIPDITFFFFY VILLAIQGAIDAFGERDQGOVKEDMETKC FICIGNDYFPDTYPFFFYHTDPITFFYV VILLAIQGAIDAFGERDQGOVKEDMETKC FICIGNDYFPDTYPFFFTHTH NTHNYDEFTHTGGOSYVWKMYQERCWE FFFAGDCFFRQYEGQ	}						i · · · · · · · · · · · · · · · · · · ·
NHILOGKLEDPAIRWOMALYKOLPNITODTS DPEKTYERVLOLANVI-FILEGKRYGKRIY CLVEHPORSKKAVWHKLLSKORKRAVVACF RMAPLYNLPRIRAVNLFLQGYEKSWIETEH YFEDKLIEDLAKFGAEPEEDEGTKRVDFLHQ LILLIFSRTALTEKCKLEEDFLYMAYADMAKS CHDEEDDDCEEVSFFEREKEKKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIALINGGNSTVQCKMLDVLKEKKDVGF FOSLAGLIMGSCSVLDLAMFERQNKAEGLGM VTEEGSGEKVLQDDEFTCJEFTLOLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEGGGRNFSKAHQVA VGFLHYFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVFKKPOFFHKAMESIKHTYQSETFFLL SCAETDENETLDYEEFVRFHEPAKDIGFNVA VLLTINLSHMPMDTRLQTFLASVLNYNGP FLGRIEMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVNNEGGEKERMELFVNFCDD TIFFMQLAAQISESDLAIERSANKESSEKERFEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKMTVKDMYAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPDPYTQBFVAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPDPYTQBFVAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPDPYTQBFVAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPDPYTQBFVAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPDPYTQBFVAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKKVAELLANMPPTTQBFVFFYBAFA LEFGGLYTTEPFFYBARGMLAFFY AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAVHYVLESSGFMEPTVRIL PLHTVISFFCIIGYYCLKYPLVIFKREKEVARK LEFGGLYTTEPFFFYBARMMLALFY AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAVHYVLESSGFMEPTVRIL PLHTVISFFCIIGYYCLKYPLVIFKREKEVARK LEFGGLYTTEPFFFYDNYBMFRAFFKFFY NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWDKFVKRKVMDWRLVINTQSFF NNYWSEDGDRAPACKYEKFROSSLSA VLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNIPFFAAHLDIAMGFFTLRTILTSVTH NGKGULTUTQLIAVVYLTYVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFIBMYVGV RAGGGGDEBDPAGDEVSTRIPDITFFFFV VILLAIIQGLIIDAFGELRDQGEVKEDMETKC FICIGNDYFDTYPHFFETHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN FFPAGDCFRKQYEDQLNC FFPAGDCFRKQYEDQLNC SEPRAGDCFRCSLAML SEPRAGDCFRCSSLAML SEPRAGDCFRCSSLAML SEPRAGDCFRCS							
DPEKTVERVLDIANVLFHLEGKSKRVORRATVACE CLMEHORSKKAVWHKLISKORKAVVACE CLMEHORSKKAVWHKLISKORKAVVACE RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH YFEDKLIEDLAKPGAEPFEDEGTKRVDPJHQ LILLFSRTALTEKCKLEEDPLYMAYADIMAKS CHDEEDDDGEEVKSFFEKEMEKQKLLVQQ ARLHDBGAAGMVQTISASKGETGFMVAAT LKLGIAILNGGNSTVQOKMLDVLKEKKDVGF FQSLAGILMQSCSVLDLMAFERONKAEGIGM VTEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQMYLRTGGNSTVTVDYLLRVQ ESISDFYWYYSGKDVIDEGGGRNFSKARIQVA KQVPNTLTEYJCQPCTGNOGSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYIQSETEFLL SCAETDEMETLDVEEFVKRFEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYVQP FIGGRIEMGSAKRIENYYFEISESTQMEKPQ VKESKRQHFDVVNEGGEKEKMELFVNFCED TIFFMQLAQIGSSDLNERSARIQWAFRSSVNSS LSKALKAMKKVKAMTVKAMTAFSSSVNSS LSKALKAMKKVKAMTVKAMTAFSSSVNSS FMTLLHFVASVFRGFFRIGSLLLGGSLVEGA KKIKVALLANMPDPTQDEVRGDGEEGRKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGGYKLIPHNPNAGLSDLSDIFGLDLKR EGGGYKLIPHNPNAGLSDLSDIFGLDLKR EGGGYKLIPHNPNAGLSDLSDIFGLDLKR EGGGYKLINTYRANFYNMIMLALFV AFAINFILLFYKVSTSSVVEGKELFTRSSSENA KVTSLDSSSHRIAVHYVLEESGYMEPTVEIL PILHTVISFCCIIGYYCLKVPLVIFKRREKEVARK LEFDGLYTTEGPSEDDLKGQMVLNTYTOSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDICAALDFSDAREKKRRKDSSLSAVLNSID VKYQMWKLIGVYTDNSFLYLAWYMTMSVL GHYNNIFFFAHLLDLAMGFKTLRTLLSSVTH NGKQLVLTVGLLAVVVLYTVVAFNFFRKF YNKSEDGOTPDMKCDDMLTCYMFHMYVGV RAGGIGGEIEDPFAGPEUTPIFFFV VULLAINGGLIDAFGELRDQCGVKEDMETKC FICIGNDYFPDFFFFTDTFFFFV VILLAINGGLIDAFGELRDQCGVVEDMETKC FICIGNDYFPDFFFTDTFFFTV VILLAINGGLIDAFGELRDQCGVVEDMETKC FICIGNDYFPDFFFTDTFFFTV VILLAINGGLIDAFGELRDQCGVVEDMETKC FICIGNDYFPDFFFFTDTFFFTFT FINYLINKDETEHTGGOSSYVWKMYQERCWE FFPAGDCFRKQYEDQLLX SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRWPPHPAHRCFCRDSLAMVL SLPRSWDYRRW							
CLVEHPQRSKKAVWHKLLSKQRKRAVVACE RMAPLYNLPHRAWNI, LDGYPKESWETEEH YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ LILLISRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEVKSFEEKMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGMVAAT LKLGIAILINGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAGLGM VTEGSGEEK VLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVMIIISTVDYLLRVQ ESISDFYWYYSCKDVIDEQORPSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELKELMDLQ KDMYVMLLSMEGRVVNOTIGKQMVDMLV ESSNNVEMILKFEDMFLKLKDLTSSDTFKEYD PDGKGVFKRPDHKAMESSTFKEYD PDGKGVFKRPDHKAMESISTFKEYD PDGKGVFKRPDHKAMESISTFKEYD PDGKGVFKRPDHKAMESISTFKEYD PDGKGVFKRPDHKAMESISTFKEYD PDGKGVFKRPDHKAMESISTFKEYD PLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYTEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFQP FLGREIMGSAKRIRRVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKEFTFEABLLSIGNDF VKYENAMFPHANALALFV AFAINFILLIFVASVFRSTWINGFTPFVIVBL FLMYLIHVASVFRGFRIGSLLIGGDEG KEEKAKEDKGKQKLRQHTHTRYGEPEPVPESA FWKIIAYQQKLLNYFANFYNMRMLALFV AFAINFILLIFYRVSTSSVEGKELPTRYSSENA KVTSLDSSRRINAHYYVLFKVEKEVARR LEFDGLYTITEQPSEDDIRGOMELVINTQSFP NNYMDKFVKRRVMDKYGEFYGRBUSELG MDKAALDFSDAREKKRYDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLIMGFRTLRTLSSYTH NGKQLVLTVTQLLAVVVYLTVVAFNFFRKF YNKSEGDDTPDMKCDMLTCYMFHMYYGV RAGGGIGDEEDPAGDEFYRRIFDTIFFFFVIVILLAUIGGIIDAFGELRDQQEVKEMETKC FICGIGINDYFOTVPHGEFTHLQENLANNYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPRGDCTRKQVEDOLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFSYAQAVCL SLPRSWDVSRWPPHPARFICICCDESSLAML							, ,
YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ LIILFSRTA.TRECKLEEDTYMAYADIMAKS CHDEEDDDGEEVKSFERKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGER VLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYYLLRVQ ESISDFYWYSGKDVIDEQRNFSKAIQVA KQVFNTLTEVIQQFCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQSQELLKELMDLQ KDMVVMLSMLEGRVVNGTIGKQMVDMLV ESSNNVEMLKFFDMFLKLKDLTSSDTKEYD PDGKGVFKRFPHKAMESHKHYTQSETEFLL SCAETDENETLDYBEFVKRFHEFAKDIGFNVA VLLTNLSEMHMPDITRLQTELAESVLNYFQP FLGREIMGSAKRIERVFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKMELFVNFCED TIFFMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHEVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDFTQDEVRGDGEGGRKP LEAALPSEDLTDLKELTEESDLLSDINSNPVPMPEVQE KFQGQXALPHNINGALSHANNPPTGDEVGCDGE KFGRQKAKEEKEEKEFTKSEPEKAEGEDGE KEEKAKEDKGKQKLRQHTHRYQEPEVVESS FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAHYTYLETRYGPEPVESS FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAHYTYLETRYGPEPVESS FWKKILAYQQKULNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIAHYTYLETRYGPEPVESS FWRYDDSFAHRLANGFYTTYGSPF NNYWDKFYKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRYDSSLSAVINSTYN GHYNNFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYTLVTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHRMYGV RAGGGGDBEDPAGDEFTRIPDTFFFFFVI VILLAIQGLIDAFGELRDQQEVKEMEFTKF TYNSEDGDTPDMKCDDMLTCYMFHRMYGV RAGGGGDBEDPAGDESTRIPDTIFFFFFVI VILLAIQGLIDAFGELRDQQEVKEMEFTKF FPLMYLINKDETEHTGQESYVWKMYQERCWE FFPRGDCTERKQVEDOLN STEPPAGDCTRAVAPSRLTAALFSYAQAVCL SLPRSWDVYRRWPHPANFICICRDESLAML	1						
LILLESRTALTEKCKLEEDT-YMAYADIMAKS CHDEEDDDGEEVKSFEEKMEKQKLLVQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT LKLGIAILNGGNSTVQKMLDYLKEKDVGF FQSLAGLMQSCSVLDLNAFERONKAEGLGM VTEEGSGEKVLQDDEFTCDLEFFLGLCEGH NSDFQNVLRTQTONNTYNGIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVPNTLTEVIQOPCTONGSLAISRLWDAV VGFLHVFAHMQMKLSQDSSQIELKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMLKFFDMFLKLKDLTSSDTFKEYD PDGKGVFKRDFHKAMESHKHVTJGSTEFLL SCAETDENETLDYEEFYKKHFEPAKDIGFNVA VLLTNLSHIMPNDTRLQTFLELAESVLNYFQP FLOREIMIGSAKRIERVYERHEPAKDIGFNVA VLLTNLSHIMPNDTRLQTFLELAESVLNYFQP FLOREIMIGSAKRIERVYENTHETAKDILTIMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVFGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNFNAGLSDLMSNFVPMFEVQE KFQEQKAKEEEKEERKSEPEKAEGEOGE KEKAKEDKOKGKLRQHTHRYGEEVYESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILAVHYVLEESGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYTTEQFSEDDIKGQWORLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDARKKKFKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFRAHLLDLINGGQWORLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDARKKKFKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFRAHLLDLINGGVDETKRITLSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDAGDEFYRRIFDITFFFFVI VILLAIIQGLIDAFGELRDQQEVKEMDHETKC FICGIGNDYPTOTPHGFETHTLQEHNLANNYLF FLMYLINKDETEHTIQESYVWKMYQERCWE FFPRADCTREKQYEDQLN SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH SSTAPANOVCH STAPANOVCH SSTAPANOVCH SSTAPANOVCH STAPANOVCH SSTAPANOVC	1						
CHDEEDDDGEEEVKSFEEKEMEKQKILJYQQ ARLHDRGAAEMVLOTKEKKOVGF FQSLAGLMQSCSVLDLMAFERQNKAEGLGM VTEGSGEKVLQDDEFDLFRFLQLLCEGH NSDRQNYLRTQTGNTTVNIIISTVDYLLRVQ ESISDFYWYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTONQQSLAHSRLWDAV VGFLHVFAHMYMKLSQDSSQIELKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVFKRDFHKABKHYTQSETFELL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERYYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFMQLAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNIITLMRMLS LKSLKKQMKLVKKMTVKDMVTAFFSSWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKR LEAALPSEDLTDLKELTESDLLSDIFGGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQBGKAKEEKEEKETSEPKARGEOGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVYBGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSYMPETTVRIL PILHTVISFFCIGYYCLKYLVIFKREKEVARK LEFGILYITEQPSEDDIKGQWDTLVNTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKREKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYYLYTVVAFNFFRKF YNKSEGGOTPDMKCDDMLTCYMFHRWYGV RAGGGIGDEIEDPAGDEYIYRIIDTFFFFVI VILLAIQGLIDAFGELRQQEGOVKEDMETKC FICGGINDFYDTVPHGFFTHTLQEHILANVLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDOLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFFF DTVLLCLSGWSAVAPSRLTAALFSNQAVCL SLFRSWDYRRWPPHPANFCIFCRDESLAML							
ARLHDRGAAEMVLQTISASKGETGFMVAAT LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFBRQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRIQLLCEGH NSDPRONYLRTQTGNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTONQGSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKKLDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETILDYEEFVKRHEPRANDGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIEDVVNEGGEKEKMELFVNFCDD TIFFMQI.AAQISESDLNERSANKESSEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVABLLANMPDPTOEVRGDGEGEGRKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPAAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGFEVPESA FWKIILAYQQKLLNYFARNFYNNMMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYTTEQPSEDDIKGQWDDLVINTQSFP NNYWDKFVRKVWNDKYGEFYGRDRISELLG MDKAALDFSDAREKKFPKSDSSAVNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNIFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFIPMYSVL GHYNNIFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFIPMYGVL RAGGGIGDEIEDPAGDEYBVRIIEDITFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFFTHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPRALSQSAGITGYTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLFRSWDYRRWPPHANNCIFCRDESLAML	1 1						
LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDPQNYLRTQTGNNTTVNILISTVDYLLRVQ ESISDFYWYSGKDVIDEQQQRNFSSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKQVIFKRDFHKARSHKHYTQSETEFLL SCAETDENETLDYEFFVKRFHEPAKDIGFNVA VLITNLSEHMPNDTRLOFILELASSVLNYQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFHFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNSDRSANKEESSKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKLVKKMTVAMDVTAFFSSYWSI FMTLLHFVASVFRGFFRIGSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNRVPMPEVQE KFQEQKAKEEKEKEEKESPEKAEGEOGE KEEKAKEDKGRQKLRQLHTHRYGEPEVPESA FWKILAYQQKLLNYFARFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEFTVRIL PILHTVISFFCIGYYCLKVPLVIFKREKEVARK LEFGGLYTTGPSEDDIKGQWDRJUNTQSFP NNYWDKFVRKVWNDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNIFFAAHLLDIAMGFKLTRILSSVTH NGKQLVLTVGLLAVVVYYTVVAFNFFRKF YNKSEDGDTPDMKCDMGFKTRTILSSVTH NGKQLVLTVGLLAVVVYYTVVAFNFFRKF YNKSEDGDTPDMKCDMGFKTRTILSSVTH NGKQLVLTVGLLAVVVYYTVVAFNFFRKF YNKSEDGDTPDMKCDMGFKTRTILSSVTH NGKQLVLTVGFTERTQCKEDVKEDMETKC FICGIGNDFFDTVPHOFFTHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFFAGDGFFRKQYEEDUN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLFRSWDYRRWPPHANFCICCEDESLAML	1						
VTEEGSGEK/LQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVRFHEPAKDIGFNVA VLLTNLSHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFMQLAAQISESDLNERSANKEESSKEPPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGGERKP LEAALPSEDLTDLFELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEGKAKEEKEEKEETKSEPKAEGEDOGE KEEKAKEDKGKÇKLIQLHTHRYGFEPVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRILMYHVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKYPLIFKREKEVARK LEFDGLYTEOPSEDDIKGGWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRKKDSLSAVLNSID VKYQMWKLGVVFUNFLYLAWYMTMSVL GHYNNFFRAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCVMFHMYYGE RAGGGIGDEIEDPAGDEVEITRIFDTIFFFFVI VILLAIQGLIDAFGELRDQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVKMYYQERCWE FFRAGGOFRKQFGDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGYTRTPSLFFP DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDVRRWYPRPHPANFCIFCRDESLAML	1						,
NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMYVMLLSMLEGNVVNGTIGKQMYDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYBEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQITLELAESVLNYFQP FLGREIMGSAKREVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFMQLAAQISESDLNERSANKESESKEREPE QGPRMAFFSILTVSALFALRYNLTLMRMLS LKSLKKQMKKVKMTVKDMYTAFFSSYWSI FMTLLHFVASVFRFICISLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEGERKP LEAALPSEDLTDLKELTEESDLSDIFGLDLKR EGGQVXLIPHNFNAGLSDLMSNPYPMPEVQE KFQEQKAKEEKEEKEEKSFEKAEGEDGE KEEKAKEDKGKQKURQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKLPTRSSSENA KVTSLDSSSHRILAVHYVLEESSGYMPETVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYVLYTVARNFFRKF YNKSEDGTPDMKCDDMLTCVMFHMYYGEV RAGGGIGDEIEDPAGDEVEIYRIFDTTFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFFAGGOFRKYGEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGYTRTPSLFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDVRRWPTPPANFCIFCRDESLAML							FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
ESISDFYWYYSGRDVIDEGGORNESKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQJELLKELMDLQ KDMVVMLLSMLEGNVVOGTIGKQMVDMLV ESSNNVEMILKFPDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFNVA VLLITNLSEHMFDTRUQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKROFIFDVNNEGGEKEKMELFVNFCED TIFFEMQLAAQISESDLNERSANKEESEKERPEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKMTVKDMVTAFFSSYWS KKIKVAELLANMPDFTQDEVRGDGEEGERKP EGGGYKLIFHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEKETKSEPEKAEGEDGE KEEKAAEDKGKQLKQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANFILLFVKSTSSVEGKELFTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYTTEQFSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRPKKDSSLSAVLNSID VKYQMWKLGVYTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLAMGFKTLRTILSSVTH NGKQUVLTVGLLAVVVYLYTVVARNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBIEDPAGDEYEIYRIFDTTFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTIGGESYVWKMYQERCWE FFFAGDCFRKQYEDQLN 446 1796 A 3592 1 3555 AGELLNSDDPPALASQSAGITGVTRTPSLFFP DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIPCRDE/SLAML							, , ,
KQVFNTLTEYIQQFCTGNQQSLAHSRLWDAV VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFPDMFLKLKDLTSSDTFKEYD PDGKGVIFKRAPKINGENSHVHYTQSETFEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQIFILELAESVLNYFQP FLGREIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFMQI.AAQISESDLNERSANKEESEKEPEEE QGFRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKKVAELLANMPDPTQDEVRGDGEEGERRP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYLCKPVLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDLIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFFP DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCICRDESSLAML]]						
VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTYQSETEFLL SCAEATDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQI. AAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVPRGFFRIGSLLLIGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEFEKSEPEKAEGEGGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNIFYNMRMLALFV AFANFILLFYSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITIGPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGGFYGRDRISELLG MDKAALDFSDAREKKPKKKDSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNIFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDPAGDEVEIYRIIFDTIFFFFVI VILLAIIQGLIAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDPAGDEVEIYRIIFDTIFFFFFVI VILLAIIQGLIAVVYVLYTVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDPAGDEVEIYRIIFDTIFFFFFVI VILLAIIQGLIAVVYVLYTVAFNFFRKF FPRAGDCFFRKQYEDQLN 446 1796 A 3592 1 355 AGGELLNSDDPPALASQSAGITGVTRTPSLFFF DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLAML							KOVENTI.TEYIOGPCTGNOOSLAHSRI.WDAV
ESSNNVEMILKFFDMELKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDYEEFVKRFHEPAKDIGFNVA VLLTNI.SEHMPHDTRI.QTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFFMQI.AAQISESDLNERSANKESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEETKSEPFEKAEGEDGE KEEKAKEDKGKULRQLHTHRYGEPEVPESA FWKKILAYQQKLLNYFARNFYNMRMLALFV AFANNFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIJAVHYVLEESGGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRWMDKYGGFYGRDRISSLLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTLLSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDIPDMKCDDMLTCYMFHMYVGV RAGGGIGDEPAGDEVEIVRIIFDITFFFVI VILLAIIQGLIIDAFGELRQQEQVKEDMETKC FICGIGNDYFDTVPHGFTHTLQEHNLANYLF FLMYLNKDETPHTGGESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPPHPANFCIFCRDESLAML]						
PDGKGVIFKRDFHKAMESHKHYTQSETEFLL SCAETDENETLDVEEFVKRFHEPAKDIGFIVVA VLLTNLSEHMPNDTRLQTIFLELAESVLNYFOP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQI.AAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNLITLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIAYQQKLLNYFARNFYNMRMLALFV AFANFILLFYKVSTSSVVEGKELPTRSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYIFTEQPSEDDIKGQWDRLVNTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTLSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYPDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLAML	1			ŀ			KDMVVMLLSMLEGNVVNGTIGKQMVDMLV
SCAETDENETLDYBEFVKRFHEPAKDIGFNVA VLLTNISEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYPEISESSRTQWEKPQ VKESKRQFIFDVYNEGGEKEKMELFVNPCED TIFFMQI.AAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERRP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQVKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEGKAKEEKKEKEFKSFEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIIEDITFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHOFETHTLQEHNLANVILF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLAMML							
VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIEN YFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFYNFCED TIFEMQI.AAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNLTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHIFVASVYFGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERRP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKILAYQKKLINYPARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFPGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVRRKVMDKYGEFYGRDRISELLG MDKAAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESSLAML						j	
FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIIFDVVNEGGEKEKMELFYNFCED TIFFMQI.AAQISESDI.NERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTI.MRMLS LKSLKKQMKKVKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKV AELLANMPDPTQDEVRGDGEEGERRP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPINPNAGL.SDLMSNPVPMPEVQE KFQEQKAKEEEKEEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFIILLFYKYSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGL YITEQPSEDDIKGQWDRL.VINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVILLCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDESLAML							•
TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKK VKKMTVKDMVTAFFSSYWSI FMTILHFVASVFRGFFRIICSILLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLINYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDK YGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEVEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVILCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				ĺ	j	j	·
QGPRMAFFSILTVRSALFALRYNILTLMRMLS LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI FMTLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGYYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKEEKEEKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIQGLIIDAFGELRDQEEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML							-
LKSLKKQMKKVKMTVKDMVTAFFSSYWSI FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGBEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEINLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQVEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML							•
FMTLLHFVASVFRGFFRICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLIGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML							
KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETK SEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKRPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML			(·			
EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEKEKETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML					ŀ		
KFQEQKAKEEEKEETKSEPEKAEGEDGE KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKDSSLSAVLNSID VKYQMWKLIGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRWPPHPANFCIFCRDE/SLA/ML			- (
KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1		1	1			•
FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML							
AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				1	İ		
PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1 1	1	}	1	-		AFAINFILLFYKVSTSSVVEGKELPTRSSSENA
LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		ł	.	1			
NNYWDKFVKRKVMDKYGEFYGRDRISELLG MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML					İ		
MDKAALDFSDAREKKKPKDSSLSAVLNSID VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	}		1	}	ł		
VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHY\NNFFFAAHLLDIAMGFKTLRTILSSVTH NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		}	ļ	-	i		· · · · · · · · · · · · · · · · · · ·
NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML							•
YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML		}	ļ		j		· •
RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML				1	ļ)	• 1
VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	}		}	}			
FICGIGNDYFDTVPHGFETHTLQEHNLANYLF FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELINSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML			Ì	1	1		
FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN 446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	}		ļ	1			
446 1796 A 3592 1 355 AGLELLNSDDPPALASQSAGITGVTRTPSLFF* DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	j j		j	1			FLMYLINKDETEHTGQESYVWKMYQERCWE
DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	1	1700		2500		355	
SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML	446	1796	A	3592	1	355	
1 1 1 1 1 1	[]		{	ı			` '
PRLVSNSWTQAILLPRPPKMLGLQV		1	- 1	- 1	j		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV
						GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI NGTLALGLKP**AWGWGEWRPKG
448	1798	Α	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM RITNENFVDA YENSNSTEFVSLASKVKDALKL LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE FSIPQHLVEEAERVMAEERVVMLPPRARSLKS FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVITVYNTL SPMEPHALVQLCGTYPPSYNLTFHSISIQNVL LITLITNTERRHPGIFEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHVKVRFKFFYLLEPGVPAGTCPKDYVEING EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPIAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVYTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV
449	1799	A	3618	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE
450	1800	A	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE HRETYQKLLEDIAVHRLAARLSSRAEVVGA

SEQ ID SEQ ID Met SEQ Predicted Predicted end Amin	o acid sequence (A=Alanine C=Cysteine,
	partic Acid, E=Glutamic Acid,
	nylalanine, G=Glycine, H=Histidine,
eotide seq- USSN location corresponding I=Isol	eucine, K=Lysine, L=Leucine,
seq- uence 09/496 correspondi to last amino M=M	ethionine, N=Asparagine, P=Proline,
	utamine, R=Arginine, S=Serine,
	reonine, V=Valine, W=Tryptophan,
	rosine, X=Unknown, *=Stop codon,
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	sible nucleotide deletion, \=possible
	otide insertion
	EKRMSKATEVMMQYVENLKRTYEKDH
	MEFKKLANQNSSRSCGPSEDGVLRTARS
	LGKNMPRRRVSVAVVPKFNALNLPGQ
	SSIPSLPALSESPNGKGSLPVTSALPALLE
	NGDPDCEASAPALTLSCLEELSQETKA
	EEAYSKGFQEGLKKTKELQDLKEEEEEQ
i I I I I I I I I I I I I I I I I I I I	PEEPEEVEETEEEEKDPRSSKLEELVHFL
	YPKLCQHWQVIWMMAAVMLVLTVVL
	NSYNSCAEQADGPLGRSTCSAAQKDSW
	GLQHEQPTEQ
	HQTVHTGRKLYECKECGKAFNQGSTLI
	UHTGEKPYECKVCGKAFRVSSQLKQHQ
	GERPYQCKELKGRGAEMLAVLAVKEQ
	VNYGK .HSAKAFHY*SSCSFSCEEGFALIGPEVV
	LGVWTAPAPVCIAVQCQHLEALNEGT
	DYPFTAFAYGSSCKYECHTVYRVRGLD BRGCYLWNGHFTT*EAISCEPLERPCH*S
	FSCEEGFALIGPEVVOCTALGVWTAPAP
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	VOCOHLEALNEGTMG
	GLGIWHVPNKSPMOHWR\KGSLLRYRT
	LOTLGHNLLGIYOKYPVKYGEGKCWT
	PVIPVVYDFGDAQKTASYYSPYGQREFT
	OFRVFNNERAANALCAGMRVTGCNTE
	GGGGYFPEASPQQCGDFSGFDWSGYGT
	YSSSREITE\AAVLLFYR
	IPAMLGLDELGRSGCGHCTQADLRFGD
	RDPGQDNDRNTAEPAFPPPPRVMAAAA
	PAQSSVTFEDVAVNFSLEEWSLLNEAQ
GCLY	HDVMLETLTLISSLGKVLILNCDLS
455 1805 A 3646 2 414 AAAC	GRGASGALTGEGGGEQGRRVGLGSRAH
SLLL	GPTFNSCQVSSQPPRVAGLGLPLKHEPS
	PSPRGPRTVRAGVPGAHPQDTPCPEFVR
	PLVGEAPGLPPEERSRGWRRDTPGLQE
	APSYDDIT
	FNSYLTLYTKNNLKSMKDLNVNTEMIK
	KNIHNLG*AKFFLN*IQKALIKRKILIHW
	IK/SFCSLSDTIKKMKRQTIVWEQTFIIHI
	LVSRIYEAFLQFNKTVNRPVFDIKKEQK
F SPACE	LOOMED WALLEY ALL SOON . TO
	LGGPTGMDLWQLLLTLALAGSSDAFSG
	AAILSRAPWSLQSVNPGLKTNSSKEPKF
	SPERETFSCHWTDEVHHGTKNLGPIQLF
	NTQEWTQEWKECPDYVSAGENSCYFN
	SIWIPYCIKLTSNGGTVDEKCFSVDEIVQ
	IALNWTLLNVSLTGIHADIQVRWEAPRN
	KGWMVLEYELQYKEVNETKWKMMDP
	SVPVYSLKVDKEYEVRVRSKQRNSGNY
GEFS	EVLYVTLPQMSQFTCEEDFYFPWLLIIF
	LTVMLFVFLFSKQQRIKMLILPPVPVPKI
	PDLLKEGKLEEVNTILAIHDSYKPEFHS
וון ון אוווין וון ווווין וווין וויין וויין	VVEFIELDIDEPDEKTEESDTDRLLSSDH
	INLGVKDGDSGRTSCCEPDILETDFNAH
DIHE	INLGVKDGDSGRTSCCEPDILETDFNAH GTSEVAQPQRLKGEADLLCLDQKNQNN
DIHE	IINLGVKDGDSGRTSCCEPDILETDFNAH GTSEVAQPQRLKGEADLLCLDQKNQNN DACPATQQPSVIQAEKNKPQPLPTEGAE
DIHE SPYH STHQ	IINLGVKDGDSGRTSCCEPDILETDFNAH GTSEVAQPQRLKGEADLLCLDQKNQNN DACPATQQPSVIQAEKNKPQPLPTEGAE AAHIQLSNPSSLSNIDFYAQVSDITPAGS
DIHE SPYH STHQ VVLS	IINLGVKDGDSGRTSCCEPDILETDFNAH GTSEVAQPQRLKGEADLLCLDQKNQNN DACPATQQPSVIQAEKNKPQPLPTEGAE AAHIQLSNPSSLSNIDFYAQVSDITPAGS PGQKNKAGMSQCDMHPEMVSLCQENF
DIHE SPYH STHQ VVLS LMD1	IINLGVKDGDSGRTSCCEPDILETDFNAH GTSEVAQPQRLKGEADLLCLDQKNQNN DACPATQQPSVIQAEKNKPQPLPTEGAE AAHIQLSNPSSLSNIDFYAQVSDITPAGS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS SCGYVSTDQLNKIMP
458	1808	A	3663	154	462	TRAPASGRSGAGLALSANAPDSGGHPGATEG PAGSLAHASGSARGTWRVRGRGSHGWERTV GAGGCANPVPALHSCASAPRGTGRVSALGPK TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC \NQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM VLALNVVTVATITVRHFVNQRADYQ\YQKLQ NY
460	1810	Α	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\ TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N CYYD/STKSFFYISCG*K\RKPTWAE\RRL\NA KMFGIPLHS\NSDPWGYEEREVIGFHRS\RV\S\RG HGS
463	1813	Α	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRIRPKSSTK
464	1814	A	3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT FSTPSSPAPFPTSSNPSSATTPPNPSP\GQR\DSR FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR WHGEVAIRLLEMDGHNQDHLKLFKKEVMN YRQTRHENVVLFMGACMNPPHLAIITSFCKG RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF \GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR EMTPGKDEDQLPFSKAADVYAFGTVWYELQ ARDWPLKNQAAEASIWQIGSGEGMKRVLTS VSLGKEVSENLSACWAFDLQERPSFSLLMD MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR FERFGLGVLESSNPKM
465	1815	Α	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP QAPLLAVGISFGGILVLNHLAQARQAAGLVA ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSYÆ*DLQARTIRQFDERYTSVA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI
466	1816	A	3684	3	307	GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE GLPDLRALLPSEDRNS SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML TGVLQG
467	1817	A	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAQLVRLAGGHCRRDTLWKRLFILE PPGPDRLRLGGRLALAELEELLEAVHAKSIGD IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFFVQPQDSESPPAQLVSTY HHLESVINTACFTLWTRLL*GSGLDH*MSLFL ESWAYQIACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ TPPRAPLPESCPLVPLTTVSHLCPLSLRVFTSHL DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQORN
468	1818	A	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA SQNGGITGMSPCAWPFFFFFFGLC
469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP HISSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFILKYLLKTQENVYNIIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEPSKEQKLIKILGDIGERVKSASDHQRQEVL KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH DACSYFTSNALPLKITFINANLMGKNISIIFKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Prolinc, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIIYRCLSTGKDQRLVQMVPDAVTLAKIHRH
470	1000		2719	430		SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGGIKNPQHFQDFVELCCRAYNIIR KHSQLLLINLLEMMLYAGILPELSGIQDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSIW*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW
470	1820	Α	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVOLLLSLGSVFLWALL
471	1821	Α	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	Α	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753		. 5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFYTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP\ERAALETPIIQGQDGSPELLIRSLV GGPSAELLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide		/=possible nucleotide deletion, \=possible
	<u> </u>			sequence		nucleotide insertion
						ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ
i						QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG
	_					LELATTFEHFYQHYMADRLLSFGSSWLEGAV
		1				LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF
						QLQRLDKLFLEQEDEEEKRL*EEEEEEEEEA
1		ŀ				EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP
j						RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG
1						PHRRLQWTWLGRAELQFGKQILHVSTVQMW
Ì						LLLKFNQTEEVSVETLLKDSDLSPELLLQALV
1						PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
						RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
1						LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL
						GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
						DVPFCGSQSETSKPSPEAVATLASLQLPAGRT
						MSPQEVEGLMKQTVRQVQETLNLEPDVAQH
						LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
						HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC
[MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD
						CPAOPTGAFIRAIVSSPEVISKYEKALLRGYVE
					,	SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
						CGWASCFNCSFPEAHYPASCGHMSQWVDDG
1						GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
						KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
						KDYYNCSAMVSKAARQEKRFQDYNERCTFH
						HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
						CQGLEQARKVLAYACVYSFYSQDAEYMDVV
•						EQQTENLELHTNALQILLEETLLRCRDLASSL
}						RLLRADCLSTGMELLRRIQERLLAILQHSAQD
{					·	FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP
						EAEEEEDDEDDVPEWQQDEFDEELDNDSFS
						YDESENLDQETFFFGDEEEDEDEAYD
475	1825	Α	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ
						ATGRRRRTRTQQRTAALLTDGTTKTGAAW
		l				SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN
						PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP
	•					DGTR\RPASITGVAQSPATRATPSLPCLHVPAP
ľ			ľ	l		SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
						RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE
1				ļ		PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS
						HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
				ļ		PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK
7,0	1020	^	3/30	701	121	RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
					1	FTMLARMVSIS*PRDLPALASQSAGITGVSHH
		į	1	ļ		APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG
ا ''' ا	1027	^	3,01	נדט	313	ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI
[]			1			RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	$\overline{\mathbf{A}}$	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
'''	1020	.^	3703	201	1240	PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS
1 1		i	i	ļ	ł	PTTSWS/PSGHSKLEVERAOTGPFCLHIYCP*P
		l	ł	!	İ	GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
			ŀ			WTGHSFASQAWLRQVPEVSKHLQCPSAESLL
	- 1	l	ľ		ŀ	TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG
			İ			PDEHPQDTDARDADGEAREREP/RRPSFAA*P
}		J			Į	VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA
			j	,]		CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE
		ļ				QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA
<u> </u>			!			Z Zoo. J. o. O. I. L. D. DOU I L. I. A.

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRFILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF VANFIMAKSWKQPTHPSVRT
481	1831	A _.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	A	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	I	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP A*FFVFI.VE\QGFTMLARMVSIS*PQ/CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	Α	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	Α	3811	378	98	RYD*SSQSENIPAQKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLLYHLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771		FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

CEO ID	CEOID	Mot	Leco	I n 1:-1-3	I mara a a a	I A A - Alo-io- C- C
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	İ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
;		ŀ	İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
488	1838	Α	3818	1	781	FRACLLELIPYAPTLSWTACPPAMAGPRGLLP
		1			1	LCLLAFCLAGFSFVRGQVLFKGCDVKTTFVT
						HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR
		l	j		!	YEVQLGGSMVSMSGCRRKCRKQVVQKACCP
		j			1	GYWGSRCHECPGGAETPCNGHGTCLDGMDR
		•				NGTCVCQENFRGSACQECQDPNRFGPDCQSV
	i					CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
						QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1839	A	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP
407	1039	^	3022	934	009	ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
			l			ODGLDLLNLMIHPPRPPKVLGFOA
490	1840	A	3825	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
.50	.0.0	1	1 3023	''	77.10	GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
						ASRPEASGDCRAGRETAMATLEKLMKAFESL
		ĺ	i			KSFQQQQQQQQQQQQQQQQQQQQQQQ
			ĺ			PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPPPPPP
			1			GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
		ŀ	:	ļ		ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
		Ì	1	1		ESDVRMVADECLNKVIKALMDSNLPRLQLEL
						YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
						CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
	·					RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
			}			KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
						TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
						LOTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
						AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
			i	İ	!	ESRSDVSSSALTASVKDEISGELAASSGVSTPG
					·	SAGHDIITEQPRSQHTLQADSVDLASCDLTSS
						ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
						TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
						GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
						FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
						VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
	,		·			APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
1						SVKALALSCVGAAVALHPESFFSKLYKVPLD
						TTEYPEEQYVSDILNYIDHGDPQVRGATALLC
						GTLICSILSRSRFHVGDWMGTTRTLTGNTFSL ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
						LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
1						LKLOERVLNNVVIHLLGDEDPRVRHVAAASL
						IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
j						KLLMHETOPPSHFSVSTITRIYRGYNLLPSITD
						VTMENNLSRVIAAVSHELITSTTRALTFGCCE
						ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
ļ						KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
į						ILAGNLLAASAPKSLRSSWASEEEANPAATK
ļ						QEEVWPALGDRALVPMVEQLFSHLLKVINIC
						AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
j						GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
ļ						GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
}						NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
1						ATLODIGKCVEEILGYLKSCFSREPMMATVC
1						
						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
				į		VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRVVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTRIQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKRENIATTHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSHSVWLGNSITPLREEEWDEEEEEE ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMMKAVAEPVSRLLESTL RSSHLPSRVGALHGVLYVLECDLLDDTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS TEXTSTURING AND AND AND AND AND AND AND AND AND AND
						RSSHLPSRVGALHGVLYVLECDLLDDTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEQLSRLDAESL VKLSVDRVNVHSPHRAMAALGLMLTCMYT GKEKVSPGRTSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA TWSLSCFFVSASTSPWVAAILPHVISRMGKLE
						QVDVNLFCLVATDFYRHQIEEELDRRAFQSV LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE FQSEWTAVV/P/EFTATQSEVADWFKDMQVP SVPIQQFPTEDWSI*PTMNDWSATSTAQTTE WVRITTEWP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK KFKRS*EKAHIRYKIDQPEDIPLKIEFLCKHSK CTATLSMRNMSLMKKKCSFSEEFLAFFPSLL VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG FTVLARMVLIS*PCDPPTLASQGTAITGMSYH ARPQDIDFLYAHQGRCWFRLL
495	1845	A	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL WADKYRPKPRFFNRVHTGFEWNKYNQTHY DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL EACADNKDFAILRFHAGPPYEDIAFKIVNREW EYSHRHGFRCQFANGIFQLWFHFKRYRYRR* RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC HGELRRHWDRLA*GPDATEGALGASFEHEG GQQPPADLTVQADTLHRPSARLGGAHRACPK RRPHRVLWRWARGAWAWRCQAREKQETQG QPCHITGHPLGREAEPAAAGAAPALAHRPPF ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE KGEGLPPGIPPGLLTGPW\SMRPVTPSFAHIR TVAPSHSSPFSGQEGRGPHGCHSPGR\SGP\AGR LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK SFVLMELAYWQDRMFF
496	1846	Α	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG LKLLTSSALPALASQSAEITGMSHRIWPLPLLR RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	A	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL
498	1848	Α	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	Α .	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGGHRTLLYGHAILRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan.
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			i	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion APDPKALRLGVLKKKAMLHQEGHMDDALSL
						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
						GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
	i . I					EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
						VNLLYELLASLIRGNRSNCALFSTNLDWLVS
	ĺ					KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
1						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
			.			ATHLRVGWALTEGYTPYPGAGEGWGGNGV
						GDDLYSYGFDGLHLWTGHVARPVTSPGQHL LAPEDVISCCLDLSVPSISFRINGCPVOGVFESF
	Í i					NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
						LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
]					LERIREKLAENIHELWALTRIEQGWTYGPVRD DNKRLHPCLVDFHSLPEPERNYNLQMSGETL
						KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE NGHNVWARDRVGQGWSYSAVODIPARRNPR
					}	LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
			1			QSGRWYFEFEAVTTGEMRVGWARPELRPDV
	1					ELGADELAYVFNGHRGQRWHLGSEPFGRPW QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
						ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
						VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
						LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
						CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
					1	NI.RRSAGGWSEAENGKEGTAKEGAPGGTPQ AGGEAQPARAENEKDATTEKNKKRGFLFKA
1					}	KKVAMMTQPPATPTLPRLPHDVVPADNRDD
						PEILNTTTYYYSVRVFAGQEPSCVWAGWVT
						PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
						VIGCLVDLATGLMTFTANGKESNTFFQVEPN
						TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
						AMFQSERKNPAPQCPPRLEMQMLMPVSWSR MPNHFLQVETRRAGERLGWAVQCQEPLTMM
						ALHIPEENRCMDILELSERLDLQRFHSHTLRL
			j			YRAVCALGNNRVAHALCSHVDQAQLLHALE
	}]			DAHLPGPLRAGYYDLLISIHLESACRSRRSML SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
1						GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
					·	ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
		l	-			RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKOILKMIEPEVFTEEEEEEDEEEGEEEDEE
	į	[}	EKEEDEETAQEKEDEEKEEEAAEGEKEEG
[į	1	ļ	ļ		LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
		İ	İ			HRVESLAAFAERYVDKLQANQRSRYGLLIKA FSMTAAETARRTREFRSPPOEOINMLLOFKDG
			ļ			TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
]]		j	j		1	GEEEEPEEETTLGSRLMSLLEKVRLVKKKEEK
] !					1	PEEERSAEESKPRSLQELVSHMVVRWAQEDF VQSPELVRAMFSLLHRQYDGLGELLRALPRA
	ļ		İ		1	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
		ļ	-		!	ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
						TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR: EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TTAKVLGSGGHASNKEKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTYM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTILI QHAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIEI ITLPMLCSYJFRWWERGPEAPPSAL PAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRAAPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMYGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAVVACFRMTIPLYNL THRACNMFLESYKAAWILTEDHISFEDRMIDD LSKAGEQEEEEEVEKKPPPLHQLVLHFSKS ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEK WMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINHICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSQIELLKEL LDLQKDMVYMLLSLLEGNVVNGMARQMV	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
contide cont	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence marce	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence 914 mg to first amino acid residue of peptide peptide peptide sequence T-Thronin, V-Valine, V-Piryptophan, Y-Tyrosine, V-Valine, V-Yilon, V-Yil	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino acid residue of opeptide sequence of pep	seq-	uence				i :	
residue of peptide sequence y=Tyrosine, X=Unknown, y=Siop codon, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide deletion, y=possible nucleoide nucle	uence			914	-		
Ppossible nucleotide disection Prossible nucleotide insertion							
sequence		1	ŀ			sequence	1 - 1 - 1 - 1
CYFCRISRONGRMFDHLSYLLENGGILGM QGSTPLDWAAASUDNINELALQEDLEX VSYLAGCGLOSCPMLVAKGYPDIGWEDGG ERYLDH RRAPVFVORGSVEPENANVVRLIR KPECFGPALRGEGSGLIAAIEARRISEDPAA DGPGIRRDRREIFGEFPERENVILGHAMM PYAALIDLIGCAPEMHLIOAGKGEALRIAL LRSLVPLEDLVGIISLPLQIPTLGKDGALVGPF MSASFVPDHKASMVLEDRVYGENOGELLEN VLDVGRLPDMRAAASLDTATSTITEMALAV NRYLCLAVLPLITKGAPLAGKGEALRIAL TRILVPDHRAAASLDTATSTITEMALAV NRYLCLAVLPLITKGAPLAGGEARMPLKLI TRILVPGROWN TO STANDARD TO		İ				1	
GGSTPLDVAAASVDDNELALALQEQDLEKY VSYLAGGGLQSCPML VAKGYPDIGWKPCGG ERYLDFLRFAVFVNGESVEENANVVVRLIR KPECFOPALRGEGGESGLAAHEEARISEDPAR DGPGURRDRREHFGGEPPEERRYULGHAM PYAALIDLLGRCAPHILOGAGGGALRIRAL LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPR MSASFVPDHKASMVLFLDRVYGIENQDFLLK VLDVGLFDMRAASLDJATRSTIEMALAV NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS MLHTVYNLSRGRSLTRAQGDVIEDCLMSLC YIRPSMLQHLURRLVFDVPILNEFAKHPLKLI TRYLFRCWKYYCLTPGWANFGCYTSEELHL TRYLFWGIFDSLAHKKYDPELYRMAMFICLC AIGAGLPDVVDASYSSKAEKKATVDAEGNI DPRPVETLNVIIPEKLDSFINKFAEVTHEKWA ADKIQNNWSYGENIDEELKTHPMRTPYKTFSE KDKESTYRWPIKESLKAMMAWEWTIEKAREGE EEKTEKKTAKISGSAOTYDPREGYNPOPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELBAKGGGTHFLLVPVDTUTLAKEKARDREK QELLKFLQMNGYAVTRGLKDMELDSSSIEKK FAFGFLQQLLRWMIDSGEFAHLEAVNSGR: EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS FAFGKLQQLLRWMIDSGEFAHLEAVNSGR: EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS FAKAVLGSGGTASKKEEMITISTFCKLAALV RHRVSLFGTDAPAVVNCHILLARSLDARTVN KSGPEPVXAGIRSFFESASEDIEKNPURL KVSQARTOVVGVGGORILTYTTVALLPVLTTLI QHAGHGFGDDVLDDVQVSCYRTLCSITSLC KVSQARTOVKGVGGORILTYTTVALLPVLTTLI QHAGHGFGDDVLDDVQVSCYRTLCSITSLC KVSQARTOVKGVGGORILTYTTVALLPVLTTLI QHAGHGFGDDVLDDVQVSCYRTLCSITSLC KVSQARTOVKGVGGARNEYARMIVARI ENTITYTEKLRPALGECLARLAAAMVAFLI PQUNEYNAGCSVTTRSPRERALIGLRNSVEEB CTOIPVERLRADIGGLAESGARTVEMHVII ITLPMLCSYLTRWWERGFEAPSALPAGAPPY CTAVTSDHLNSLLGNILBLINNNICIDEASWM KRLAVFAQPIVSRARPELLQSHFPTIGRLRKR AGKVVSEEQLALEARAAGAGGIGGLLVRDEFS VLCRDLYALYPLIRYVDNNRAQWLTEPNBS AELERMWGEITFYWSSHFPREEGRIPVVQ NEINMMSFLTADMKSKMAKAGDIOSGGSODD RIKKKRRGDRYSVGTSLIVATLKKMLPIGLN MCAFTDODLITLAKTRYALKDTDEEVREFL HNILLIGGKVEGSSPERWOMALVYRLDGTEHPYKS KKAWWHLLISKORRAVAPCFRITLYNLI THRACMFLESYKAAWILTEHSFEDRMIDD LSKAGGGEEBEEVEFEKREMPHLYDL, ULFIERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVEFEKREMPHLYDL, ULFIERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVEFEKROPHLYDL, ULFIERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVEFEKROPHLYDL, ULFIERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVEFEKROPHLYDL JUFFERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVEFEKROPHLYDL JUFFERS ALLTEKSKLDEDYL JWAAYADMAKSCHLEEG GENGEGEEBEEVERNER ORDO		L			sequence		
VSYLAGGCLQSCPML/VAKGYPDIGWKFCD RYLDFLRAVVORGSVERNAVVVRLIR KPECFGPALRGEGGSGLAAIEEAIRISEPA DOPGIRDRERREHFGEPPEERVYHLGHAIM FYAALIDILGRCAPEMH-IQAGKGEALRIRA LRSL VPLEDI-VGISLPLQIPILGGALVOPE MSASSVPDHKASMVL-LDRVYGIENODPLIL VLDWOFLPDMRAAASLDTATEMALAV NRYLCLAVLPLITKCAPLPAGTEHRAIMVDS MHTVYRLSGRSLTKAQRDVDCLMSLCG YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLI TNHYBRCWKYYCLPTGWANFGVTSEELHI TRKLFWGIFDSLAHKKYDPLYIMAMPCLC AIAGALPTDYVDASYSKAEKKATVDAERON DPRRVETLNVIJIPEKLDSFINIK-REYTHERWA DKIQNINWSYGENIDEBLKTHPMLRPYKTSE KDKETYRWIKESIK.KAMM WEWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNIPOPD SAVTLSRELQAMAEQLAENYHITWIGRKKK ELEAKGGGTHPLLVPYDTLTAKEKADREK, QELLKFLQMMGVAYTGLKDMELDSSSIEK FAGFIQULRWMDISGEFIAHLBAVVSGE EKSPHEQEKTFAKILLPLINGYFTNHCLYFLS TPAKVLGSGGHASNEKEMITSLFCKLAALU RHRVSLFGTDAAVVNCHILARSLDARTVM KSGFEIVKAGLRSFFESASEDIEKMYENLING KYSQARTOVKGVOQNLTYTVALPYLULTLI QHIAQHOFGDDVILDDVQVSCYRTLCSIVILL KYSQARTOVKGVOQNLTYTVALPYLULTLI QHIAQHOFGDDVILDDVQVSCYRTLCSIVILL TIKNTYYEKLRPALGECLARLAAMPVAFLI PQLNEYNACSVYTTKSPREALICJENSVERE CPDIPVLERLMADIGGLAESGRYTEMPHVII ITLPMLCSYLFRWWERGFEAPPSALPAGAPPI CTAYTSDHLNSLIGHBILINGBEASWM KRLAVFAQPIVSRARPELQSHIFPITGRIKK AGKVYSEEGGLAAKAAGPOKJEPPSALPAGAPPI CTAYTSDHLNSLIGHBILINGBEASWM KRLAVFAQPIVSRAPRELLQSHIFPITGRIKK AGKVYSEEGGLALEAKARAGEGLANVSTENL MCAPTDQDLITLAKTRYALDDTEHPKIR AGKVYSEEGGLALEAKARAGEGLINDERS VLCRULYALJYPLLIRYYDNNRAQWLTEPNS AEELFRAVGEIPTWSKSINNFKRALIGELSSAWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSHIFPITGRIKKMINGENSWM KRLAVFAQPIVSRAPPELLQSH							
REYLDELRAVVNOGSVERNAVVVRLIL RECFOPALRGOGSGLLANGEARISEDPAR DOPGIRRORRENFHCGEEPPEENRY-LICHARIA PAAALIDLLGRGAEPMH-LJOAKGREALRIRA LRSLVPLEDL VOIISLP-LOIPTLGKOGAL-VOPP MSASTVPDHKASMVLFLDRVYGLOPDLLI VLDWGFLPDMRAAASLDTATESTTEMALAV NRYLCLAVIPLITKCAP-LPAGTEHRAIMVDS MLHTVYRLSRGSLTKAGRDVEDLAMSL YIRPSMLOHLURRIV-PFOVPILINEFAKMPLLC YIRPSMLOHLURRIV-PFOVPILINEFAKMPLLC AIAGAL-PPDVYDASYSKARKATVDAEGNI DPRVETLINVIIPEKLDSFINKFABYTHEW ADKLONNWSYGENDELEKTHPMR-PYKTYSE KOKENYR WPIKESLKAMM WEWTIEKAREGE EKKTEKKKTAKJIGSOGATVDPHR-PYKTYSE KOKENYR WPIKESLKAMM WEWTIEKAREGE EKKTEKKKTAKJIGSOGATVDPHR-PYKTYSE KOKENYR WPIKESLKAMM WEWTIEKAREGE EKKTEKKKTAKJIGSOGATVDPHARPYNTTWGRKKK KELEAKAGGTHIPLLVPPOTLTAKEKARDREK, GELLKFLOMIGVAYTRGLKDMELDSSIESE FAFOFI-QOLL RWMDISGEFIAHLLEAVVSSGR EKSPHEOEIKFFAKIL-LPLINQYFTINHCLYKK KSOPETVAGAURASNEKEMITIST FCKLAALV RHRVSLEGTDAPAVVNCLHILARSLDARTVY KSOPETVAGAURASNEKEMITIST FCKLAALV RHRVSLEGTDAPAVVNCLHILARSLDARTVY KSOPETVAGAURASNEKEMITIST FCKLAALV RHRVSLEGTDAPAVVNCLHILARSLDARTVY KSOPETVAGAURSSFESSASDIEWWENLING KVSOARTOVKGVGONLTYTTVALLPVLTTIL QHIAQHOPGDDVILDDVOYSCYRT LCSIYSL TIRNTYYEKLRPALGECLARLAAMPVAFLI RQINEYNACSVYTTKSRERALLCLPRIVERE CPOIPVLERLMADIGGLASSARTIEMPHVII ITI.PMLCSYLPRWERGFGAPPSALPAGAPPY CTANTSDHLINSLICHMILINGHENDENSMIK KRLAVFAQPIVSRARPELLQSHIPTIGRIKKI AGKVYSEEGOLALEAKABGOGEGLLVRDES VLCRDLYALYPLLIRYYDNNRAQWLTEPNS AEELFRWYGEIPTYWSKSINFREREQONFVVQ NEINNMSFLTADINSKMAKAGDIQSGGSBOON KTKKKRRGDBYSVGYSTLINGTHILLYNLINGHILLYNDLINGEASWM KRLAVFAQPIVSRARPELLQSHIPTIGRIKH AGKVYSEEGOLALEAKABAGGEGLLVRDES VLCRDLYALYPLLIRYYDNNRAQWLTEPNS AEELFRWYGEIPTYWSKSINFREREQONFVVQ NEINNMSFLTADINSKMAKAGDIQSGGSBOON KTKKKRRGDBYSVGYSTLINGTHILYNLINGTHILYNLINGTHILYNLINGHENDLINGTHILYN							
REFEGGALAGIEANISEAN VICTORIAN PAALIDLIGRCAPEMHLIQAGKGEALRIKA LRSL VPLEDL VGIISH PLOPITLOGAL VOPE MASAS VPDHKASMVLFLDR VYGIENQDFLLH VLDVGFLPOWRAAASLDTATESTIEMALAV NRYLCLAVIPLITKCAPLPAGTEIRAIMVDS MLHTVYRLSKGRSLTKAQRD VIEDCLISKLEY VRESMLQHLERL VFDVPLINEFAKMPLKLI TINHYERCWKYYCLPTOWANFGVTSEELLH TRKLFWGIFDSLAHKKYDPELYRMAMPCLC ALAGALPPDYVDASYSSKAEKKATVDAERD DEVELOW DEVELOM DEVELOW DEVELOM DEVEL	ļ	•				ĺ	
DOFGIRRDRRRHTGEEPPERNYHLGHAIN FYAALIDILIGRACPEMHLIQAGKGEALRIRAI LRSLVPLEDLVGIISLPLQIPTLGKDGALVOPP MSASTVPDHKASMAVLELDRVGINOPLLI VLDVGFLPDMRAAASLDTATFSTTEMALAV NRYLCLAVLPLITKCAPLPAGTEHRAIMVDS MLHTVYRLSRGSLTKAQRDVGLIMSUC YIRPSMLQHLLRRLVFDVPLINEFAKMPLLCL TRKLFWGFDSLAHKKYDPELYRMAMPCLC AIAGALPPDYNDASYSKAEKATVDAEGNI DPRPVEILMVIIPEKLDSFINKFAEYTHEKWA DKIQINNWSYGGNDEELKIHPMR.PPKYTFSE KDKETRWPIKESLKAMAWEWTIEKABEGE EEKTEKKKTAKJKOSQATVDPREN SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGTHPLLVFYDTLTAKEKAPDEK QELLKFLQMINGYAVTRGLKDMLDSSSIEK FAFGFLQULLRWMDISGEFIAHLEAVVSSG EKSPHEQEKTFAKILLFLINGYFTNHCLYFIS TPAKVLGSGGHASNKEKMITSLFCKLAALU RHRVSLFGTDAAVVNCLHLASLDARTVM KSGPEIVKAGLRSFFESASEDIEKMPENILG KVSGARTOVKGVGONLTYTVALPVLITLI QHIAQHOFGDDVILDDVQVSCYRTLCSIVAL TIKNTYYEKLRALGELARAMPVAFLI PQLNEYMACSVYTTKSPREALIGLIPNSVEB CPDIPVLERLMADIGGLAESGARTTEMPHY ITILPMLCSYLFRWWRGFGAFPSALFAGAFPI CTAYTSDHINSLLGHILRINNNLGIBEASM KRLAVAQPIVSRARPELLGSHFPINGIRIKR AGKVVSEEGIALEAKABQGEGLLVTDEN KRLAVAQPIVSRARPELLGSHFPINGIRIKR AGKVVSEEGIALEAKABQGEGLLVTDEN AEELFRUNGEFIYWSKSINFKREEDIFVVQ NEINNMSFLTADINSKMAKAGDGISGGSDOI RTKKKRRGDRYSVQTLIHRYDNINKAQWITEPNE AEELFRUNGEFIYWSKSINFKREEDOFIVVQ NEINNMSFLTADINSKMAKAGDGISGGSDOI RTKKKRRGDRYSVQTLIHRYDNINKAQWITEPNE AEELFRUNGEFIYWSKSINFKREEDOFIVVQ NEINNMSFLTADINSKMAKAGDGISGGSDOI RTKKKRRGDRYSVQTLIHRYDNINKAQWITEPNE AEELFRUNGEFIYWSKSINFKREEDOFIVVQ NEINNMSFLTADINSKMAKAGDGISGGSDOI RTKKKRRGDRYSVQTSLIVATLKKMLPIGIN MCAPTDOOLTILAKTRALADIOSGGSDOI RTKKKRRGDRYSVQTSLIVATLKKMLPIGIN MCAPTDOOLTILAKTRALADIOSGGSDOI RTKKKRRGDRYSVQTSLIVATLKKMLPIGIN MCAPTDOOLTILAKTRALADIOSGGSDOI RTKKKRRGDRYSVQTSLIVATLKKMLPIGH HTRACNMFLESKYRAWULTEDHSFEDRIMG ALTEKSKLDEDVI,YMAYADIMAKSCHLEEG GENGEAEEEVEVEYSFEREMEKREQULYQQAR HTRGAABMVLQMISACKGETGAMVSSTILL GISILNGGNABVYSFEREMEKREQULYQQAR HTRGAABMVLQMISACKGETGAMVSSTILL GISILNGGNABVYSFEREMEKREQULYQULARINGGABPOPP LEGNNDFONYLRTTOTINTITTNIICTVDYLL RUGESSIDFYWYNGSGRAWAVACGFORMYSKAL SVAKQVYNSLTEYJQGCTGROOSSLAHISHLW DAVVGFLHVFAHMMMKLAQDSSOIGLLKEL		i					
FYAALIDLLGRCAPEMH_UJQAGKGEALRING LRSLVPLEDL_VGBIS_PL_QIPL_GROGAL_VQPK MSASS_PVPDHKASM_VLE_DR_VYGIEN_OPEL_LIV_ VLDVGFLPDMRAAASLDTATTEMAL_AV NRYLCLAVL_PLITKCAPL_PAGTEHR_AIM_VDS MLHTVYRLSRGRSLTK_AGRD_VIEDCLMSEL_ FYRENAL_GRILLER_VFO_VPILLER_FAKM_PLK_LI TRIK_PWGIPDS_LAHKK_VTDPL_YRMAMPCLC GYRENAL_GRLER_VFO_VPILLER_FAKM_PLK_LI TRIK_PWGIPDS_LAHKK_VTDPL_YRMAMPCLC AIAGAL_PPD_YVDASYSSKAEKK_ATVDAEGN DREVETL_NVIIP_BKL_DSFINK_PASY_THE WA DKIQN_NWSYGEN_DEEL_KTHPML_PYK_TTSE CHECKET_REVET_PAGE_GREEN_CREATER_FASY_HE REVET_PAGE_GREEN_CREATER_FASY_FASY_HE REVET_PAGE_GREEN_CREATER_FASY_FASY_HE REVET_PAGE_GREEN_CR							
ILSSLYPLEDILVGIISLPLQIPTI_GKDGALVQPK MSASPVPDHKAASMV_IENOPGIENOPLIL V1_DVGFLPDMRAAASLDTATFSTTEMALAV NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCE YIRPSMLQHLLRRILVFDVPILNEFARMPLKLI TINHYERCWKYYCLPTGWAGYTSEELHLI TIKLFWGIPDSLAHKKYDPELYRMAMPCLC AIAGALPPDVYDASYSSKAEKATVDAEGNI DPRPVETLNVIIPEKLDSFINKFASYTHEKWA DKIQNNWSYGENIDELEKLTHPMLRPYKTFSE KOKEIYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKYTAKISOSAQTVDPREGYNPQPDP SAVTLSRELQAMABQLAENYHNTWGRKKK ELEAKGGTHPLLVPYDTLTAKEKARDEK QELKFIQMNGYAVTRGLKOMELDSSIEIEK FAFGFLQQLLRWMDISQEFIAHLBAVYSGRE EEKSPHEQEIKFFAKILLPLNGTHNICLYFIS. TPAKVLGSGGHASNKEKEMITSLFCKLAALV RRRVSLFGTDAPAVNCLHLLARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMYENIRLG KVSQARTOWKGYGQNLTYTTVALLPVLTTLL QHAQHQFGDDVILDDVOYSCYRTLCSIYSLC TIKNTYVEKER,PPALGELALAAAMPVAPAL PQLNEVNACSVYTTKSPRERALLGLPNSVEER CPDIPVLERLMADIGGLAESGAYTTTVALLPVLTTLL QHAQHQFGDDVILDDLOVOYSCYRTLCSIYSLC TIKNTYVEKER,PPALGELALAAAMPVAPAL TOLIPHCSYLPRWWERGPEAPPSALTAGAPPY CTAYTSDHLNSLGNILRIVNNLCIDEASWM KRLAYFAQPIVSRAPPELLGSHIPFITGRLRKR AGKVVSEEQLALEAKAAQGGEGLLVRDERS VLCRDLYALYLLIRVVDNAQWLTEPNS AEELFRMVGEIFIYWSKSHNKREEGONFVOQ NEENMSGLTADNKSKMAAGQUISGGSGDOG RTKKKRRGDRYSVOTSLIVATLKKMLPIGIN MCAPTTOPOLITAAKTRYALRQWLTEPNS AEELFRMVGEIFIYWSKSHNKREEGONFVOQ NEENMSGLTADNKSKMAAGQUISGGSGSDOG RTKKKRRGDRYSVOTSLIVATLKKMLPIGIN MCAPTTOPOLITAAKTRYALRQWLTEPNS AEELFRMVGEIFIYWSKSHNKREEGONFVOQ NEENMSGLTAADNKSKMAAGDIOSGGSGDOG RTKKKRRGDRYSVOTSLIVATLKKMLPIGIN MCAPTTOPOLITAAKTRYALRQWLTEPNS AEELFRMVGEIFIYWSKSHNKREEDOMPVOQ NEENMSGLTAADNKSKMAAGDIOSGGSGDOG RTKKKRRGDRYSVOTSLIVATLKKMLPIGIN MCAPTTOPOLITAKTRYALRQWLTEPNS ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGBAEEEVEKKPROPLHQLVLHPSSI ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGBAEEEVEVSFEEKQMEKQRLLYQQAR HTGAAEMVLQMISACKGETGAMVSTILLI GISINGGNAECKGTGAMVSTILLI GISINGGNAECKGACHTGMYSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGTGAMVSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI GISINGGNAECKGACHTGRAWSTILLI G		ļ					i l
MSASTVPDHKASMVLFILDRYYGIENODPLIA VILDYGIEDDMRAASLIDTSTTEMALAY NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS MHITVYRLSRGRSLTKAQRDVIEDCLMSLCR YRRSMLQHLLRRLVPDVPINEFAKMPLKLI THKREWGIPSLAHKKYPDHINEFAKMPLKLI THKREWGIPSLAHKKYPDELYRMAMPCLC AIAGALPPDVDASYSSKAEKKATVDABGNI DPRYPETLNVIIPEKLDSFINAFAYTHEKWA DKIQNNWSYGENIDEELKTHPMLRPYKTFSE KDKETYRWPKESLKAMLAVETISKE KDKETYRWPKESLKAMLAVETISKE KDKETYRWPKESLKAMLAVITAKEKKE ELEAKGGGTHPLLVPYDTITAKEKARDREK, QELLKFLQMMGYAVTRGLKDMELDSSIEKF FAFGFLQQLLRWDMISOGFIAHLEAVVSSGR ESPHEQEIKFFAKILLPLNQYFFINICLYFLS TPAKVLGSGGIASNKEKEMTISLFCKLAVL RRVSLFOTDAPAVVNCLHLLARSLDARTVN KSGPEIVKAGLRSFFEASEGHMYENLRJ KVSQARTQVKGVGQNLTYTTVALLPVLTTLI QHAQHQFODDVILDVQVSCYRTLCSYSLC TTKNTYVEKLPALDEJOVASCATICLSYSLC TTKNTYVEKLPALDEJOVASCATICLSYSLC TTKNTYVEKLPALDGGLAEARAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLENSVEEN CPDIPVLERIMADIGGLAEARAAMPVAFLI TITLPMLCSYLPRWWERGFEAPPSALPAGAPPI CTAVTSDHINSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFPITGLRKR AGKVVSEEGIALEAKAAGOEGELLVNDESS VLCRDLYALPYLIRYVDNNRAQWLTEPNPS ABELFRANGEIFIY WSKSHNRAEEGIPFV VQ NEINMSSLTADNKSKMAKAGDIQSGGSDGJ RTKKKRRODRYSVGTSLIVATLKKMLPIGLN MCAPITODOLITLAKTRYALTEDHSFEDRMIDD LSKAGEGEEEEVEEKKPDPLHQLVLHFSR ALTEKSKLDEDYLYMAYADIMAKSCHLEG GENGBAEEEVSVSFEKKQMEKGRLLYQQAR HTGAACMVLOHISSYKAAWLTEDHSFEDRMIDD LSKAGEGEEEEVEEKKPDPLHQLVLHFSR ALTEKSKLDEDYLYMAYADIMAKSCHLEG GENGBAEEVSVSFEKKQMEKGRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSTLKL GISINGGNAEVQNKLMLYLDGTHFULGUL GGTNINGNGEKVMADDEFURFLUNL GISINGGNAEVGVSFERKGMEKGRLLYQQAR HTGAACMVLOHISACKGETGAMVSTLLL GISINGGNAEVGVSFERKGMEKGRLLYQQAR HTGAACMVLANJARGROPGREHYSALM SVAKQYFNSLTEVIQGPCTGRQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW DAVVGFLHVFAHMMKLAQDSSGLAHSRLW	İ						,
VILDVGFLPDMRAAASLDTATFSTTEMALAY NRYLCLAVLPLTKCAPLPAGTHRAMYDS MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCF YIRPSMLQHLLRRLVPDVPILNEFAKRHPLKLI TINHYERCWKYYCLPTGWANFGYTSEELHLI TRKLFWGIPDSLAHKKYDPELYRMAMPCLC AIAGALPPDVYDASYSSKAKKATVDAEGNI DPRPVETLNVIIPEKLDSFINKFAEYTHEKWA DKIQNINWSY GENIDELELKTMLRYPYTFEE KDKETYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNRQPPD SAVTLSRELQAMAEQLAENHINTWGRKKKK ELEAKGGGTHPLLVPYDTLTAKEKARDEK, QELLKFLQMNGYAVTRGLMBLDSSSIEKF FAFGELQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFAKILFLYNGWLDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEIKFAKILFIAHLEAVYSGR EKSPHEGEICHALTAKARALGELAALAWPWELL GINAGMAWALTTYTVALLFYTHILARLEAVYSGR EKSPHEGEICHALTAKARALGENSEEN CHAMPAGPUSCARTEN KARARTON TOTT TOTT TOTT TOTT TOTT TOTT TOTT	1	1					
NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS MLHTYPRISCRISTIKAQDIDECLMSLC YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLI TINHYBERCWKYYCLPTGWANFGOTISEELHL TRKLFWGIFDSLAHKKYDELYRMAMPCLC AIAGALPPDYVDASYSSKAEKKATVDABGNI DPRPVETLNVIIPEKLDSFINAFAYTHER WA DKIQNNWSYGENIDEELKTHPMLRPYKTSE KOKEIYRWPIKESLKAMLAWEWTIEKABEGE EKEKKKATAKISQSAQTYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTLTAKEKAPREKK QELLKFLQMNGYAVTRGLKDMELDSSSEKT FAFGFLQQLLRWMDISQEFMISLEVCKYKK QELKFLQMNGYAVTRGLKDMELDSSSEKT FAFGFLQQLLRWMDISQEFMISLFCKLYFL AND RIFWSLFOTDAPAVYNCLHILARSLDARTVM KSGPEIVKAGIASFFEASGGMSNKEREMTSLFCKLANL RRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGIASFFEASGGMSNEREMMISLFCKLAST KYSQARTQVKGVGQNLTYTTVALLPVLTTL QHIAQHQFGDDVILDDVQVSCYTLSYLSC TITKNTYVFKLRPALGECLARLAAMPVALTL PQLNSYNACSVYTIKSFRERAILGLENSVEEN CPDIPVLERIMADIGGLAESGARYTEMPHVII TILPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSHILNSLLGNILRIVALIGIBASWM KRIAVFAQPIVSRARPELLQSHIPTIGRIKR AGKVYSEEGLALEAKAAQGGELLVRDEFS VLCROLYALLYPLLRYVDNNRAQWLTEPNSP AEELFRMVGEIPTYWSKSHNFKREQNIPVVQ NEINNMSTLTADNKSKMAKAGDIQSGGSDGJ RTKKKRRGBRYSVGTSLIVATLKKHLIPGIN MCAPITODOLITALKTRYALNDTDEEVREFLH NICHLGGKVEGSPSLRWQMALYRGVPGREE DADDPEKVRRVQEVSTALYLDTGHEPYRS KRAVWHKLLSKQRRRAVVACRRMITYNIL THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGGQEEEEEVEEKKPDPLHQLVLHFSR; ALTEKSKLIDEDYLYMAYADIMAKSCILEEG GENGBAEEEVSVFEEKKWERGRLYQQAR HTRGAAEMVLQMISACKGETGAMVYSTLLAL GISINGGNAEVQNKLMLYLDGIFRHQUL GISINGGNAEVMADERSYLAVGTRMTYNIL THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGGQEEEEEVEKKPDPLHQLVLHFSR; ALTEKSKLIDEDYLYMAYADIMAKSCILLEG GGRGAEEEVSVFEEKKWERGRLYLQQAR HTRGAAEMVLQMISACKGETGAMVSTLLL GISINGGNAEVMADERSYLAVGTRMTYNIL THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGGGEEEEEVERKPOPLHQLVLHFSR; ALTEKSKLIDEDYLYMAYADIMAKSCRILEG GGRGAEEEVSVFEEKKWERGRLYLQQAR HTRGAAEMVLQMISACKGETGAMVSTLLL GISINGGNAEVMADGREKVMADGREKVMADGREKVMADGREKVMADGREKVMADGREKVMADGREKVMADGREKNTSKAM SVAKQVFNSLTEVIQGPCTGRQQSLAHSRLW DAVVGFIHVFAHMMMKLAQDSSGJELKEL LDLQKMVVYNGMLSLLEGWKYNGMARQMVY	İ	Ì					•
MLHTVYRLSRGRSLTKAQROVIEDCLMSLC YIRPSMLQHLLRRLVPDVPIINEFARMPLKLI TNRYERCWKYYCLPTGWANFGVTSEELIH TRKLFWGIFDSLAHKKYDPELYRMAMFCLC AIAGALPPYVDASYSSKAEKKATVDAEGNI DPRPVETLNVIIPEKLDSFINKFAKTYTHEK WA DKIQNINWSYGENIDEELKTHPMLRPYKTTSE KDKETYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNPOPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTLTAKEKARDREK, QELLKELQMINGYAVTRGLKDMELDSSIEKS FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEGEIKFFAKILLNVYSTINHCLYFLS TPAKVLGSGGHASNKEKEMITSLECKLAALV RHBYSLFGTDAPAVVNCHIILARSLDARTVW KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALPVLTTLL QHAQRQFGDDVILDDVQVSCYRTLCSIYSL TTKNTYVEKLRPALGECLARLAAMPVAFLI PQLNEYNAGCSVTTKSPRERAILGLRNSVEE CPDIPVLERLMADIGGLAESGARYEMPHVII TLPHLCSYLPRWWERGFEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHIPTIGRLRIK AGKVVSEEQLALEAKAEAQEGELLVRDERS VLCRDLYALYPLLIRYVDNNRAQWILTEPINS AEELFRMVCEIFIYWSKSHMFKREEONPVVQ NEINNMSFLTADNKSKMAKAGDIOGSGSDD RTKKKRRGDRSVSVGTATLKMHLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFIH NNLHLQGKVEGSPSLRWQMALYROVGERPHYS KKAVWHKLLSKQRRRAVVACFRMTIYNIL THRACNMFLESYKAAWULTEDISFEDNIDE LSKAGEQEEEEEVEKKPDPLHAYLTHAKHLPIGLN KKAAWHKLLSKQRRRAVVACFRMTIYNIL THRACNMFLESYKAAWULTEDISFEDNIDE LSKAGEQEEEEEVEKKPDPLHAYLTHAKHLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFIH NNLHLQGKVEGSPSLRWQMALYROVGEREFIH NNLHLQGKVEGSPSLRWQMALYROVGEREFIH NNLHLQGKVEGSPSLRWQMALYROVGEREFIH NNLHLQGKVEGSPSLRWQMALYROVGEREFIH GEGNGEAEEEVEVSFEKQMEKQRLLYQARR HTRGAAEMVLQMISACKGETGGANVSTIKL GISILNGGRAEVQCKMLDYLLKKKEGFFQ IQALMQTCSVLDLNAFERQNKAGGLGMVVBE DGTVINROMGEKVMADDEFTOLFRFLQLL GISILNGGNAEVQCKMLDYLLKKKEGFFQ IQALMQTCSVLDLNAFERQNKAGGLGMVVBE DGTVINROMGEKVMADDEFTOLFRFLQLL GGINNDFONYLRTQTGNTTTINITUTVYLL RLOESISDFVWYSGKAGLGGMVVBE DGTVNTROMGEKVMADDEFTOLFRFLQLL EGHNDFONYLRTQTGNTTTINITUTVYLL RLOESISDFVWYSGKAGLGGMVVBE DGTVNTROMGEKVMADDEFTOLFRFLQLL EGHNDFONYLRTQTGNTTTINITUTVYLL RLOESISDFVWYSGKALELEGOGMVNSTIKL DAVVGFLHFVFAHMMMKLAQDSGGLLKERL DAVVGFLHFVFAHMMMKLAQDSGGLLKERL	1	i					
YIRPSMLQHLLRRLJFDVPILISFAMPIKLILI TNYFYERCWKYYCLPTGWANFGYTSEELLILI TRKLFWGIFDSLAHKKYDPELYRMAMPCLC AIAGALPPDYVDASYSSKAEKKATVDAEGNI DPRYPETLINVIIPEKLIDSFINKFAEYTHEKWA DKIQNNWSYGENIDEELKTHPMLRPYKTTSE KDKEYTRWPIKESIKAMAWEWTIEK AREGE EEKTEKKKTAKISQSAQTYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTLTAKEKARDREK QELLKFLQMNGYAVTRGLKDMEDSSIEISK FAFGFLQOLLRWMDISGEFIAHLEAVYSSGR EKSPHEQEIKFFAKILLPLNQYFTHICLYFIS TRAKVLGSGGHASNKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGKRSFFESASEDIEKMFSLRCLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGKRSFFESASEDIEKMFSLRCLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMFSLRCLAALV TKNTYVEKLRPALGECLARLAAMPVALI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSL TTKNTYVEKLRPALGECLARLAAMPVALI PQLNEVNACSVYTTKSPRERALIGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVII ITLEMLCSYLPRWWERGEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNICIBEASWM KRLAVFAGPIVSKARPELQSHFPTIGRIRKR AGKVVSEEGOLALEAKAEAQEGLLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNES AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIOSGGSDDI RTKKKRGDRYSVGTAUTAKMLPIGLN MCAPTDQDLITLAKTRYALKDTIECEVREFILI NNIHLQGKVFGSPSLEXWQMALYRGVPGREE DADDPEKIVRRVGEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLI THRACKMFLESYKAA WULTEDHSFEDRMIDL LSKAGEQEEEEEVEEKRPDPLHQLVLHFSRS ALTENSKLDEDYLYMAYADIMAKSCHLEG GENGEAEEVEVSFEKQMEKGRLLYQQAX HTTRAAEMFLESYKAA WULTEDHSFEDRMIDL LSKAGEQEEEEEVEEKRPDPLHQLVLHFSRS ALTENSKLDEDYLYMAYADIMAKSCHLEG GENGEAEEVEVSFEKGMEKGRLLYQQAX HTTRAAELDEDYLYMAYADIMAKSCHLEG GENGEAEEVEVSFEKGMEKGRLLYQQAX HTTRAAEMFLESYKAA WULTEDHSFEDRMIDI LGKAGEQUEEEEGEVEKRPDPLHQLUL BEGINNDFQNYLRTQTGTNTTTINITUTDYLL RICCESSEDYFWYSGKORDIECKGRINSKAM SVAKQVFNSLTEYIQGPCTGNQGSLAHSRLW DAVVGFLHVFAHMMMRLAQDESQIELLKRLU DAVVGFLHVFAHMMMRLAQDESQIELLKRLU DAVVGFLHVFAHMMMRLAQDESQIELLKRLU DAVVGFLHFVAHMMMRLAQDESQIELLKRLU DAVVGFLHFFAHMMMRLAQDESQIELLKRLU DAVVGFLHFFAHMMMRLAQDESQIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAHMMMRLAQDSGIELLKRLU DAVVGFLHFFAH							
TNIYERCWKYYCLPTGWANFGOYTSEELLH TRKLFWGIFDSLAHKKYDPELYRMAMPCLC AIAGALPPDYVDASYSSKAEKKATVDAEGNI DPRPVETLNVIIPEKLDSFINKFAETHEKWA ADKIQNNWSYGENIDEELKTHPMLRPYKTYSE KDKEIYRWPIKESLKAMIAWBWTHEKAREGE EEKTEKKKTAKISQSAGTYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTILTAKEKARDREK, QELLKFLQNNGYAVTRGLKDMELDSSSIEKT FAFGFLQQLLRWMDISQEFIAHLEAVYSSGR: EKSPHEQEKFFAKILLYNYFTHICLYFLS TPAKVLGSGGHASNKEKEMITSIFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSOPEIVKAGLRSFFESASEDIERWHENLRIG KVSQARTQVKGVGQNLTYTTVALLPVLTTL QHAQHQFGDDVLIDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAMPVAFLI PQLNEVNAGSVYTTKSPRALIGLINSVEEN CPDIPVLERLMADIGGLAESGARYEMPHVII TILTHALCSYLPRWWERGPBAPPSALPAGAPP CTAVTSDHLNSLLGNILRIIVNNLGIBCASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLKK AGKVVSEEQLALEAKARAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPINES AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADINKSKMAKAGDIOSGGSDDJ RTKKKRGDRYSVGTATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFIH NNLHLQGKVFGSPSLRWQMALYTROPS TKKKRGDRYSVGTATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFIH NNLHLQGKVFGSPSLRWQMALYTROPS LSKAGGQEEEEEVVEKRPDPHOLVLHFINS KKAVWHKLLSKQRRRAVVACFRMTPLYNLL THRACNMFLESYKAAWULTEDHSPEDRINDI LSKAGGQEEEEEVVEKRPDPHOLVLHFINS ALTENSKL DEDN'L YMAYADIMAKSCHLEEG GENGEAEEEVVENFEKOMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGARMVSTIKL GISILNGGNAEVQQKMLSCKGFTGS IQALMQTCSVLDLNAFERONKAEGLGMYVBE DGTYNTROMGES WMADDEFTOLHFRILCILDL GGINNDFONYLRTQTGTNTTTINICTUPYLL RIGGSILNGGNAEVQGKMUSTELKEL GISILNGGNAEVQGKMUSTELKEL GISILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVGCKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKMUSTLICTUPYLL RIGGSILNGGNAEVQGKORUSSTLIKU DAVVGFLHYFAHMMMKLAQDSGELLKEL LDLUKDMVVMLLSLILEGNVVNGMIARQMV	1	1					
TRKLFWGIPDSLAHKKYDPELYRMAMPCLC AIAGALPPDYVDASYSSKAEKKATVDAEGNI DPRPVSTLNVIIPEKLUSSPINKFAEYTHEKWA DKIQNNWSYGENIDEELKTIPHALPYKTFSE KDREIYRWPIKESIKAMAWBWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNPQPPD SANTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTLTAKEKARDREK, QELLKFLQNINGYAVTRGLKDMERKKK ELEAKGGGTHPLLVPYDTLTAKEKARDREK, QELLKFLQNINGYAVTRGLKDMELDSSSIEKT FAAFGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEQEIKFFAKILLPLNQYFINHCLYFIS, TPAKVLGSGGHASNKEKMITISLFCKLAALV RHEVSLFGTDAPAVVNCLHILARSLDARTVN KSGPEIVKAGKGGGNASTKEKMITISLFCKLAALV RHEVSLFGTDAPAVVNCLHILARSLDARTVN KSGPEIVKAGKGVGONLTYTTVALLPVLTTL QHAQHQFGDDVILDDVQVSCYTTLCSIYSIC TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERALIGPISVERSH TIKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERALIGPISVERSH CTAVTSDHLNSLLGNILRIINNIGIEASWM KRLAVFAQPIVSRARPELQSHFIFTIGRLRK AGKVVSEEGQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLRYVALYPALGYONRAQWLTEPNS AGELFRMVGEIFIYWSKSHNFKRECONFVVQ NEINNMSCHTADNKSKMAKAGJOSGOSGOSIO RTKKKRRGDRYSVGTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFIL NNLHLQGKVFGSPSLENAVLTKMPLIVINL THRACMMFLESYKAA WULTEDHSFEDRMID LSKAGEQEEEEEVEEKKPPDLHQLVLHFSKT ALTENSKLDEDYLYMAYADIMAKSCHLEGG GENGEAEEVEVSFEKKMEKGRLYQQAM HTTRACAMFLESYKAA WULTEDHSFEDRMID LSKAGEQEEEEEVEEKKPPDLHQLVLHFSKT ALTENSKLDEDYLYMAYADIMAKSCHLEGG GENGEAEEVEVSFEKGMEKGRLLYQQAM HTTRACAMFLESYKAA WULTEDHSFEDRMID LSKAGEQEEEEEVEEKKEPDLHQLVLHFSKT ALTENSKLDEDYLYMAYADIMAKSCHLEGG GENGEAEEVEVSFEKGMEKGRLLYQQAM HTTRACAMFLESYKAA WULTEDHSFEDRMID LGKAGEQEEEEGEVEKKPPDLHQLVLHFSKT ALTENSKLDEDYLYMAYADIMAKSCHLEGG GENGEAEEVEVSFEKGMEKGRLLYQQAM HTTRACAMFLESYKAA WULTEDHSFEDRMID LGKAGEGEGEGGGFOMYSSTLAL GISILNGONAEVQQKMLDYLKDKKEVGFFO; IQALMQTCSVLDLNAFERQNACEGGGMYMB DGTYNTRONGEEV WANDDEFTQLFRFIQLL EGHNNDFQNYLRTQTGNTTTINITUTDYLL RQESISDFYWYTYGGKGTGGGGGMYRB DGTYNTRONGEEVANDADEFTQLFRFIQLL EGHNNDFQNYLRTQTGNTTTINITUTDYLL RQESISDFYWYTYGGKOWMADLEFQOKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHFVFAHMMMKLAQDSGIELLKEL		1					
AIAGALPPDYVDASYSSKAEKKATVDAEGMI DPRVETLNVIIPEKLDSYNKFAEYTHEKWA DKIQNNWSYGENIOEELKTHPMLRPYKTFSE KDKETYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKTAKISGSAQTYDPREGYPNOPPD SAVTLSRELQAMAEQLAENYTHNTWGRKKK ELEAKGGTHPLLVPYDTLTAKEKARDEK, QELLKFLQMNGYAVTRGLKDMELDSSSIEKT FARGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEGEIKFFAKILPLNQYFTHICLYFLS TPAKVLGSGGHASNKEKEMITSI-FCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVW KSGFEIVKAGLRSFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLL QHAQHOFGDDVILOVOYSCYRTLCSIYSLC TTKNTYVEKLRPALGECLARLAAAMPVAFLL PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPSALPSAGAPPP CTAYTSTDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALFYPLIKTYDNNRAQWLTEPNES AEELFRNVGEIFIYWSKSHNFKREEQNFVVQ NENNMSFLTADINKSKMAKAGDIOSGGSDOJ RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTTRYALKDTDEEVREFLH NNHLQGKVEGSPSLRWQMALYKGVPGREE DADDPEKIVRRVQEVSAULYYLDQTEHPYKS KKAVWHKLLSKORRRAVVACFRMTPLYNLL THRACNMFLESYKAAWILTEDHSPEDRMIDL LSKAGEQEEEEEVEEKKPDPLHQLVLHFSR ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMERGRILLYQQAR HTRGAAEMVLQNISACKGETGAMVSSTLKL GISILNGGMAEVQOKMLDYLKDKKEVGFFG IQALMOTCSVLDLNAFERQNKAEGLIGMYNE DGTVINRQNGEK VMADDEETQDLFRFLQLLC EGHNDRONYURTQTGTNTTTINIICTVDYLL RLOESISDFYWYYSGKDVIEGOGKRYNSTLMA SVAKQVFNSLTEYJQGPCTGNQGSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL		1					
DPRPVETLNVIJPEKLDSFINKFAEYTHEK WA DKIQNNWSYGENIDEELKTHPMLRPYKTFSE KDKEYYRWIKESLKAMIA WEWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNPQPPD SAYTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGTHPLLVPYDTLTAKEKARDEK, QELLKFLQMNGYAVTRGLKDMEDKSSIEKI FARGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEQEIKFFAKILLPLINQYTNHCLYFIS. TPAKULGSGGHASKEMITISTFCKLALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KYSQARTQVKGYGQNLTYTTVALLPVLTTL QHIAQHOFGDDVILDDVQVSCYRTLCSIVSLC TTKNTYVEKLRPALGECLARLAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVII ITLPMLCSVLPRWENGFEAPPSALPAGAPP CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHIPTIGRIRIR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLRYYDNNRAQWLTEPNES VLCRDLYALYPLLRYYDNNRAQWLTEPNES AELFRMYGEIFTYWSKSHNFKREEONFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDIDEEVREFLH NNLHLQGK VEGSPSLRWQMALYRGVFGREE DADDPEKLVRRVQEVSRAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLL THRACNMFLESYKAAWLLTEDHSFEDRMIDDL LSKAGGQEEEEEVEKKPPDHQIU VLHFSR: ALTEKSKLDEDYL YMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKKMPDHJU VLHFSR: ALTEKSKLDEDYL YMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKKMERDLRIJCHLING GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMQTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMGTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMGTCSVLDLMSACKGETGAMVSSTLKL GISILNGGNAEVQOKMLDYLKDKREVGFFG IQALMGTCSVLDLMSACKGETGAMVSSTLKL GLOSSDFYWYYGGROTITINIICTVDYLL RUDGSSDFYWYYGGROTIONQSLAHSRLW DAVVGFILHYFAHMMMKLAQDSSQIELLKEL LDLQKOMVYMLLSLLEGNVVNGMARQMY					:		
DKIQNNWSYGENIDEELKTHPMLRPYKTTSE KDKETYRWPIKESLKAMMAWEWTIEKAREGE EEKTEKKKTAKISQSAQTYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGTHPILLYPDTUTLAKEKARDREK, QELLKFLQMNGYAVTRGIKDMELDSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEQEIKFFAKILPLINQYFTHHCLYFLS TPAKVLGSGGHASNKEKEMITSLFCKLAALV RHRVSLEGTDAPAVCKGVGQNLTYTTALLEVLUTTL QHAQHQFGDDVILDDVQVSCYRTLCSIYSLC TTKNTYVEKLRPALGECLARLAAMPVAFLI PQLNEYNACSVYTTKSPREALIGLPRSYEEN CPDIPVLERLMADIGGLAESGARYTEMPHVII TILPMLCSYLPRWWERGPEAPSLPAGAFP CTAYTSTDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHIPTIGRLRKR AGKVVSEEQLALEAKARAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFTYWSKSINFKEEQNFVVQ MENNNMSFLTADNKSKMAKAGDIOSGGSDQ RTKKKRRGDRYSVOTSLIVATLKKMLPIGIN MCAPTDQDLTTLAKTRYALKTDEEVREFLH NNLHLQKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAALVYLDGTEHPYKS KKAVWHKLLSKGRRAVVACFRMTILYNL THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEOEEDEEVEKKPPPLHQLVLHFSR ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVVSFEEKQMEKQRLLYQQAR HTRGAAGAWVLQMIXSTLKL GISILNGGNAEVQKMLDYLKDKKEVGFFQ IQALMQTCSVLDLNAFERQNKAEGLIGNVSSTLKL GISILNGGNAEVQKMLDYLKDKKEVGFFQ IQALMGTCSVLDLNAFERQNKAEGLIGNVNE DGTYNRQNGEKVMADDEFTQDLTFRFLQLLC EGHNDROPLYTLRTQTNTTINIICTVDYLL RLOESISDFYWYYSGKDVIEGQKKRYRSTAM SVAKQVFNSLTEYJQGPCTGNQGSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL							
KDKEIYRWPIKESLKAMIAWEWTIEKAREGE EEKTEKKKTAKISQAQIYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKK ELEAKGGGTHPLLVPYDTITAKEKARDREK, QELLKFLQMMQYAVTRGLKDMELDSSIEISK FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR' EKSPHEQEIKFFAKILLPLNQYFTIHCLYFLS TPAKVLGSGGHASKEKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGYGONLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPIVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRAPPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNES AEELFRMVGEIPTWSSHNFKREEQNFVVQ NEINMSFLTADNKSKMAKAGDIOSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALVRGVPGREE DADDPEKLVRRVQEVSAVLYVLDOTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDMIDDL LSKAGEQEEEEEEVEKKPPPLHQLVLHFRSK KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDMIDDL LSKAGEQEEEEEEVEKKPPPLHQLVLHFRSK ALTEKSKLDEDVLYMAYADIMAKSCHLEGG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDARFERQNKAEGLGMYNGE DGTVINRQNGEKVMADDEFTQDLFFFLQLLC EGHNNDFQNYLRTQTGGNTTTINNIICTUVDYLL RLQSISDFYWYYSGKDVIEEQGKRIFSKAM SVAKQYFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSQIELLKEL LDLQKDMVVNILLSLLEGNVVNOMARAQMU							_
EEKTEKKKTAKISQSAQTYDPREGYNPQPPD SAVTLSRELQAMAEQLAENYHNTWGRKKKK ELEAKGGGTHPLUPYDTLTAKEKARDREK, QELLKFLQMNGYAVTRGIKDMELDSSIEKI FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR: EKSPHEQEIKFFAKILLPLNQYFTNHCLYFLS; TPAKVLGSGGHASNKEKEMITSLFCKLAALV RHRVSLFGTDAPAVVCLHILARSLDARTYW KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLYTTVALLPVLITLJ QHIAQHQFGDDUDDVVSVSYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEEQLALEAKAAQEGELLVRDERS VLCRDLYALYPLLIRYVDNNRAQWLTEPNES VLCRDLYALYPLLIRYVDNNRAQWLTEPNES AEELFRWVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSTLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGIN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKILSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAA WILTEDHSFEDRMIDL LSKAGEQEEEEEVEEKKPDPLHQLVLHEPST ALTEKSKLDEDYLYMAYADIMAKSCHLEG GENGGAEESVEFEKQMEKGRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDJAFFRQNKAEGLGMYNE DGTYNRQNGEX YMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTNIIICTVDYLL RLQESISDFVWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKLAAQDSSQIELLKEL LDLQKDMVVNGLILSLEGNVVNGMARQMY	ļ						
SAVTLSRELQAMAĒQLĀENYINTWGRKKK ELEAKGGTHPLLVPYDTLTAKEKARDREK, QĒLLKFLQMNGYAVTRGLKDMELDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR' EKSPHĒQĒRIFĀKILLPLNQYFTHHCLYFLS TPĀKVLGSGGHASNKEKEMĪTSLFCKLĀĀLV RHRVSLFGTDĀPĀVVNCLHLARSLDARTVM KSĢESIVKĀGLRSFFESASEDIEKMVENLRLG KVSQĀRTQVKGVGQNLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG TIKNTYVEKLRFALGĒCLĀRLĀAAMPVĀFLI PQLNEYNĀCSVYTTKSPRĒRĀILGĻPNSVĒEN CPDIPVLĒRIMĀGIGLĀESGARYTEMPHVĒ ITLPMLCSYLPRWWERGPĒAPPSĀLPĀGĀPPI CTĀVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLĀVPĀQPIVSRĀRPĒLLQSHĒPTIĞRLRIK ĀGKVVSĒĒQLĀLĒAKĀRĀQĒĢĒLLVRDĒFS VLCRDLYĀLYPLLIRYVDNNRĀQWLTĒPNPS AĒĒLFRMVGĒLĪTYWSKSHNFKRĒĒQNFVVQ NĒINNMSFLTĀDNKSKMĀKĀGDĪGSGGSDQI RTKKRRĀGDR ŠVQOTSLIVĀTLKKMĻPĪĢLM MCAPTDQDLĪTLĀKTRYĀLKDTDĒEVRĒFĒL NNLHLQGKVĒGSPSLR WQMALYRGVPĒRĒĒ DĀDDPĒKIVRRVĢEVSĀVLYVLDOŢĒHPYKS KKĀVWHKILSKQRRRAVVĀCFRMĪPLYNLĒ THRĀCNMFLĒSYKAĀWILTĒDHSFĒDRMIDĒ LSKĀGĒĢĒĒEEVEVĒKĶPDPĻHQL VLHFRST ĀLĪĒKSKLDĒDVI, YMĀVADĪMAKSCHLĒGĒ GĒNGĒĀĒĒVĒVSFĒEKQMĒKQRLLYQQĀR HTRGĀĀEMVLQMISĀCGĒTGĀMVSSTILKL GISILNGGNĀEVQQKMLDYLKDKKEVGFFĢS IQĀLMQTCSVLDAFFROŅKAĒGLGMVNĒ DĞTVINRQNĒEK VMADDĒFTQDLFFFLQLLC ĒĞHNNDFĢNYLRTQTGTMTTTINIICTUVĪLL RLĢSISDFYWYYSGKDVIĒEGGKRNFSKĀM SVAKQYFNSLTĒPIQĢPCTGNQOSLAHSRLW DĀVVĢFLHVĀHMMMKLAQDSSQIĒLLKĒL LDLQKDMVVNIKLISLĒGONVNOMĀRQMV							
ELEAKGGTHPLLVPYDTLTAKEKARDREK. QELKFLQMNGYAVTRGLKDMELDSSSIEKF FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR EKSPHEQEIKFFAKILLPLNQYFTNHCLYFIS. TPAKVLGSGGHASNKEKEMITSLFCKLAALV RIRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENLRIG KVSQARTQVKGVGQNLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLC TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERALIGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVII ITLPMLCSYLFRWWERGPEAPPSALPAGAPPI CTAVTSDHLINSLLONILRIIVNNLGIDEASSWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALJYLLIRYYDNNRAQWLTEPNES AEELFRMVGEIFTYWSKSHNFKREEQNFVVQ NEINMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTISLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTILYNLL THRACNMFLESYKAAWLTEDHSFERRMID LSKAGEQEEEEEVEEKKPDPHQLVLHFSRT ALTEKSKLDEDYTLWAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQAR HTRGAABMVLQMISACKGETGAMVVSTLKL GISILNGGNAEVQQKMLDYLKKKKVGFFGS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTNIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVPISLTEYTQGCTGRQQSLAHSRLW DAVVGFLHVFAHMMKKLAQDSSQIELKEL DLDLQKDMVVMLLSLLEGNVVNGMARRQMV		i i					
QELLERLQMNGYAVTRGLKDMELDSSSIEKE FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR: EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TPAKVLGSGGHASNEKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVQQNLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLC TITKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEVNACSVYTTKSPRERALIGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITILPMLCSYLFRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPITOQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDL LSKAGEQEEEEEVEEKKPDPLHQLVLHFSR: ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGBAEEEVVSFEEKQMEKQRLLVQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLDKEVGFFQ IQALMQTCSVLDNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFONYLRTOGTNITTNIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGCFTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKELL DLDLQKDMVVMLLSLLEGNVVNGMARQMV							
FAFGFLQQLLRWMDISQEFIAHLEAVVSSGR: EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS TTAKVLGSGGHASNKEKEMITSLFCKLAALV RHRVSLFGTDAPAVVNCLHILARSLDARTYM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTILI QHAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIEI ITLPMLCSYJFRWWERGPEAPPSAL PAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRAAPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMYGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRAVVACFRMTIPLYNL THRACNMFLESYKAAWILTEDHISFEDRMIDD LSKAGEQEEEEEVEKKPPPLHQLVLHFSKS ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEK WMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINHICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSQIELLKEL LDLQKDMVYMLLSLLEGNVVNGMARQMV					,		QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
EKSPHEGEIKFFAKILLPLINQYFTNHCLYFIS TPAKVLGSGGHASNKEKEMITSLFCKLAALV RRHVSLFGTDAPAVVNCLHILARSLDARTVM KSGPEIVKAGLRSFFESASEDIEKMVENLRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLI QHAQHOFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERALIGLPNSVEES CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALTYPLLIRYVDNNRAQWLTEPNS AEELFRMYGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTODQLTLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPPPLHQLVLHFSRI ALTEKSKLOPTLYNAMADIMASCKILEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTILKL GISILNGGNAVQAKMLAYADIMASCKLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTILKL GISILNGGNAVQAKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEXYMADDEFTQDLFRFLQLLC EGGNNDFQNYLRTQTGNTTTINIICTVUYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYJQOPCTGNQQSLAHSRLW DAVVGFLHYFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMARQMV		!					FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
RHRVSLFGTDAPAVVNCLHILARSLDARTVN KSGPEIVKAGLRSFFESASEDIEKMVENIRLG KVSQARTQVKGVGQNLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVII ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAYTSDHINSLLGNIL RIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYAL YPLLIRVYDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTIPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDY'LYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGANVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFGG IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEPIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKKEL		i					EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
KSGPEIVKAGLRSFFESASEDIEKMVENI.RLG KVSQARTQVKGYGQNLTYTTVALLPVLTTLI QHAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEITIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNIS SKMAKAGDIOSGGSDQD RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDE LSKAGEQEEEEEVEKKRPDPLHQLVLHFSR: ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQE IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQLLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGRQQSLAHSRLW DAVVGFLHVFAHMMKLAQDSSQIELKEL							TPAKVLGSGGHASNKEKEMITSLFCKLAALV
KVSQARTQVKGVGQNLTYTTVALLPVLTTLI QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLC TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AELFRWYGEIFIY WSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAA WILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTITNIICTVDYLL RLQESISDFYWYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQGSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV		1					RHRVSLFGTDAPAVVNCLHILARSLDARTVM
QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEVNACSVYTTKSPRERALIGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVPAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNBS AEELFRMYGEIFIY WSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDITDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMITILYNLE THRACNMFLESYKAA WILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRI ALTELSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTITINIICTVDYLL RLQESISDFYWYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQGSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELKKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV						·	KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
TTKNTYVEKLRPALGECLARLAAAMPVAFLI PQLNEYNACSVYTTKSPRERAILGLPNSVEER CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMILCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDL LSKAGEQEEEEEVEEKKPDPLHQLVLHFSR: ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVVSFFEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFOS IQALMQTCSVLDLNAFERQNKAEGLGMVNS DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV		!					KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
PQLNEYNACSVYTTKSPRERAILGLPNSVEEN CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYXS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDL LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKYMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	}						, , , ,
CPDIPVLERLMADIGGLAESGARYTEMPHVIE ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRWVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDOJ RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMIPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
ITLPMLCSYLPRWWERGPEAPPSALPAGAPPI CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDDJ RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRI ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEEVESFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTTNIICTVDYLL RLQESIDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMKKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV		1					•
CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPPPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLE THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTIINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
AGKVVSEEQLALEAKAEAQEGELLVRDEFS VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAA WILTEDHSFEDRMIDE LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQAR HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNDDFQNYLRTQTGNTTITNIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTITNIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
AEELFRMVGEIFIYWSKSHNFKRÈEQNFVVQ NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKLVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFFFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							•
NEINNMSFLTADNKSKMAKAGDIQSGGSDQI RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKLYRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1						
RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFFFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV					-		
MCAPTDQDLITLAKTRYALKDTDEEVREFLH NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFTFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1						·
NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDD LSKAGEQEEEEEVEEKKPDPLHQLVLHFSST ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							
DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDE LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQ IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1	[.]					1
KKAVWHKLLSKQRRRAVVACFRMTPLYNLF THRACNMFLESYKAAWILTEDHSFEDRMIDE LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1						DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
THRACNMFLESYKAAWILTEDHSFEDRMIDE LSKAGEQEEEEEVEEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							KKAVWHKLLSKORRRAVVACFRMTPLYNLP
LSKAGEQEEEEEVEKKPDPLHQLVLHFSRT ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							THRACNMFLESYKAAWILTEDHSFEDRMIDD
ALTEKSKLDEDYLYMAYADIMAKSCHLEEG GENGEAEEEVEVSFEEKQMEKQRLLYQQARI HTRGAAEMVLQMISACKGETGAMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1						LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
HTRGAAEMVLQMISACKGETGÀMVSSTLKL GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	'						
GISILNGGNAEVQQKMLDYLKDKKEVGFFQS IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1	1					HTRGAAEMVLQMISACKGETGAMVSSTLKL
DGTVINRQNGEKVMADDEFTQDLFRFLQLLC EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV						ŀ	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL RLQESISDFYWYYSGKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	ł	1		' I			IQALMQTCSVLDLNAFERQNKAEGLGMVNE
RLQESISDFYWYYSĞKDVIEEQGKRNFSKAM SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV							DGTVINRQNGEKVMADDEFTQDLFRFLQLLC
SVÄKQVFNSLTEYIQGPCTGNQQSLAHSRLW DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV	1						` `
DAVVGFLHVFAHMMMKLAQDSSQIELLKEL LDLQKDMVVMLLSLLEGNVVNGMIARQMV]					RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
LDLQKDMVVMLLSLLEGNVVNGMIARQMV	į						
	1	1		İ			7 7
I I DMLVESSSNVEMILK FFDMFI KI K DIVGSF A							•
							DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI	L	L			·		QDY V I DPKGLISKKDFQKAMDSQKQFSGPEI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QFLLSCSEADENEMINCEEFANRFQEPARDIG FNVAVLLTNLSEHVPHDPRLHNFLELAESILE YFRPYLGRIEIMGASRRIERIYFEISETNRAQW EMPQVKESKRQFIFDVVNEGGEAEKMELFVS FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA AEAGAEGAEEGAAGLEGTAATAAAGATARV VAAAGRALRGLSYRSLRRVRRLRRLTAREA ATAVAALLWAAVTRAGAAGAGAAAGALGL LWGSLFGGGLVEGAKKVTVTELLAGMPDPT SDEVHGEQPAGPGGDADGEGASEGAGDAAE GAGDEEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG VDGVEEELPPEPEPEPEPELEPEKADAENGEK EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS GGSSGWGLGAGEEAEGDEDENMVYYFLEES TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP LVIFKREKELARKLEFDGLYTTEQPEDDDVKG QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG DIYGRERIAELLGMDLATLEITAHNERKPNPP PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
	,					AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV KEDMETKCFICGIGSDYFDTTPHGFETHTLEE HNLANYMFFLMYLINKDETEHTGQESYVWK MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	A	3869	467	665	VIVAIYCQLIFDKGAKTIQ+PFQQIAL/CKRMK LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	Α	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL L
504	1854	В	3896	279	70	MVSKSKSII.MSYNHVELTFSDMKKMPEAFRR TQKHTIYLIPYQVIFWSTGKDAMRSFMMPFY QKEYYENQ*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG NENTKLELRKVPPELNNISKLNEHFSRFGTLV NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS ASSDLPQVLSTLLA*QKQCIIQLL/WKAAQKT LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA VHGRGRGRGRGRGRGVPGHAVVDHRPRALEIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA
506	1856	A	3911	1952	919	VITFKTRAEAEAAAVHGARFKGQDLKLAWN KPVTNISAVETEEVEPDEEEQREIIIA DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	Α	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKQRDKRNRHLGR
508	1858	Α	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	Α	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	Α	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	I	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPKVTTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRMP\SP MAALILVADNAGGSHASKDANQVHSTTRRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	Ā	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGQARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863		3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- ucnec	Met hod	SEQ ID NO; in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN VELGITISSPSSRKPGDPLVISDIKKGSVAHRT GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR YGGPLGVITISGTEEPVFDL*IISSLTKGGLAERT GAIHIGDRILLAINSSSLKGKPLSEAIHLLQMAG ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSAVIDTSYGTEGTVSFQASGYVNFNTYD WRSPKQRGSVLSPVTVKPRSQTYPDVGLSYED WDRSTASGFAGAAVDSAETEQEENFWSQALE DLETCGQSGILRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFFQQPSHGGNLETREPTVIL
514	1864	Α	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC
515	1865	Α	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	Α	3977		1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARK VDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV
517	1867	Α	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	peptide	1100	in NO.	nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
nucl-		i			location	
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Ì		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	ł			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	l	t		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	'	/=possible nucleotide deletion, \=possible
Į.	l			sequence		nucleotide insertion
		<u> </u>		Sequence		LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
ŀ	1				,	
L						AHV\FADLLLIITLPSYYIPFC
520	1870	Α	3999	882	698	QSFRLSLLSSWDYRHM*PRLANF*T\FFCRDR/
		l	ĺ			SLALLPRLVSNSWPQAILPPRPPKVLGLQT
521	1871	A	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
			ŀ		j	PPTSASHVAGATGTHHHAWLLSV
522	1872	A	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
1 322	10,2	^	4013	_	1 3''	
1					ļ	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
	i	ļ				EYGPVYSTWSALEGELAEPLEGVSACIGNCST
<u> </u>					İ	AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
1	l]			[KVKEVCKTSKS/GOVIYKGVSIRLRANFLAEP
ł	1	l			1	L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
1		ł			[GNEGDITSFPAK
524	1874	-	4020	1067	743	
324	18/4	Α .	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
ŀ		ļ	[TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
1		1	[ĺ	RQGFTVLARMVLIS*PHDLPASASQSAGITGL
ł		Ì	1		}	SHCSWPTSSILS
525	1875	Α	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
1	10.0	**	.021	.01	33.	PPRLKO/F/SHLSPPSIWDYRRVPPCLVNFSIFF
j						VETGSCOPCLOLLGSSNPPASASOSAGIAGISH
1					i	
[QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
						NKIINRPTHPVESSF
526	1876	Α	4024	80	341	TPSSTSRGTEEQQSSKMAWQRREEKEHLNVR
1						RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ
		l				LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	A	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
1 32 /	,0,,	^	1020	373	250	DVAVYFTTKEWAIMG\PAERALYRDVMLEN
i I			l			
						YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
<u></u>						TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	Α	4028	1160	242	GTSELLCIQRWNWGPAFPPRPGLALAPTLQLL
						VEMGSAKSVPVTPARPPPHNKHLARVADPRS
						PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ
1					·	DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE
1						VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
[
						QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
[ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
[NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
						AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
<u> </u>						WEQGQD\HDKENQHFPLVES
529	1879	Α	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
		}				CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI,*T
]			•			ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
<u> </u>	1000			250		HTEICT*MFIAVLFVVVKTWKQF
530	1880	A	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
						RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
Į l						/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
į l						VKISILPKVNYRFYLISIKIIMAI
531	1881	Α	4061	50	278	TOGTEEIYKISSCEWVQASFSTPLITLHDFKIY
		**			_,,	HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
1						IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
						T*KR
532	1882	Α	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
						YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
!				i		KLFONTE/TEGRHPISLYEFRITLITIPNKDNIYL
ţ						QIWMPVSLMNIVTLKCPT
1522	1002		4076		255	
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/
					İ	ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
<u> </u>						KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC

nucleotide sequence Deptide sequence Deptide s	SSHWNC IFSHQVL AFDMQR IANRKQE
sequence Sequence	SSSHWNC IFSHQVL AFDMQR IANRKQE
uence 914 ng to first amino acid residue of peptide sequence geglutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan Y=Tyrosine, X=Unknown, *=Stop codor /=possible nucleotide deletion, \=possible nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS 1931 IIDSSTRRMESERSPLYRQLIDLGYLS GAPGQDTKAQSMLVEQSEKLRHLST QTRLVDAAKALNUVHCHCLDIFINQ DLQITPKRLEYTRKKENELYESLMN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	SSSHWNC IFSHQVL AFDMQR IANRKQE
amino acid residue of peptide sequence	SSHWNC IFSHQVL AFDMQR IANRKQE
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS 534 1884 A 4088 3 1931 IIDSSTRRMESERSPLYRQLIDLGYLS GAPGQDTKAQSMLVEQSEKLRHLS QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKENELYESLMN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	SSHWNC IFSHQVL AFDMQR IANRKQE IEFKDVI
sequence nucleotide insertion IFNAIPIKMPMMCMAKIEKNSS 534 1884 A 4088 3 1931 IIDSSTRRMESERSPLYRQLIDLGYLS GAPGODTKAQSMLVEQSEKLRHLS: QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKENELYESLMN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	SSSHWNC IFSHQVL AFDMQR IANRKQE IEFKDVI
534 1884 A 4088 3 1931 IIDSSTRRMESERSPLYRQLIDLGYLS GAPGQDTKAQSMLVEQSEKLRHLS: QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKENELYESLMN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	TFSHQVL AFDMQR IANRKQE IEFKDVI
534 1884 A 4088 3 1931 IIDSSTRRMESERSPLYRQLIDLGYLS GAPGQDTKAQSMLVEQSEKLRHLS QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKKENELYESLMIN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	TFSHQVL AFDMQR IANRKQE IEFKDVI
GAPGQDTKAQSMLVEQSEKLRHLST QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKKENELYESLMIN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSGS LWEQIKQIIQRITWVSPPAITLEWKRI	TFSHQVL AFDMQR IANRKQE IEFKDVI
QTRLVDAAKALNLVHCHCLDIFINQ DLQITPKRLEYTRKKENELYESLMIN EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	AFDMQR IANRKQE IEFKDVI
EMKDMIVETLNTMKEELLDDATNM VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	IEFKDVÌ
VPENGEPVGTREIKCCIRQIQELIISRI NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	
NKLISSVDYLRESFVGTLERCLQSLE VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	LNQAVA
VHITSNYLKQILNAAYHVEVTFHSG: LWEQIKQIIQRITWVSPPAITLEWKRI	MOODING
LWEQIKQIIQRITWVSPPAITLEWKRI	
ESLSASKLAKSICSQFRTRLNSSHEA	
LEAGHSGRLEKTEDLWLRVRKDHA	
LESRSLQDVLLHRKPKLGQELGRGQ	
CDNWGGHFPCALKSVVPPDEKHWN	
HYMRSLPKHERLVDLHGSVIDYNYC VLLIMERLHRDLYTGLKAGLTLETR	
VEGIRFLHSQGLVHRDIKLKNVLLDI	
TDLGFCKPEAMMSGSIVGTPIHMAP	
YDNSVDVYAFGILFWYICSGSVKLPI	EAFERCA
SKDHLWNNVRRGARPERLPVFDEE	
ACWDGDPLKRPLLGIVQPMLQGIMI	VRLCKS\
NSEQPNRGLDDST 535 1885 A 4090 2 417 ALMPHEANYEEIFLKTDKDMDGFES	GIEVRE
iFLKTR/GLPSTLLAHIWALCDSKDCG	
HFALAFHLIT\QKLIKGIDPPLVLTPEI	
ASLQKVTELTRKPVCIIFKGTILWRIT	dsiwmk
HNRKRIWLRA 536 1886 A 4102 569 829 DHOK*KNIPCSWIGRINIVKMSILPK/	ADVDECAL
536 1886 A 4102 569 829 DHQK*KNIPCSWIGRINIVKMSILPKA PIKIPMTFFTEI*S*NVYRTTKTQE*AI	
EQNLEESHYLDFK*YYRAV	MUDDICIC
537 1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRLTIAQIKE	HKWML
IEVPVQRPVLYPQEQENEPSIGEFNE	QVLRLM
HSLGIDQQKTIE	
538 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFA VPRENITAWMNAIGLIITALPVS	AGCSQPL
539 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIA	SSEI+MM
I*KFAT\KKSPGPYRFTAEFSHTFKED	
PLFPKIYREGTLPHSFYEASITL	
540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRO	
AALGDFLGLHRRTQQPAVDRLLSDA	
VRGHGGVRESGRAPQQPGRRRGRRJ GRWRREGCGAGGRGVCVAAWSQR	
DYRLFHKMSNSHPLRPFTAVGEIDH	
IGALLIGEEYGDVTFVVEKKRFPAHF	
CQYFRALLYGGMRESQPEAEIPLQD	
MLLKYIYTGRATLTDEKEEVLLDFL	
GFPELEDSTSEYLCTILNIQNVCMTFI	
SLPKLTCMCCMFMDRNAQEVLSSEG TALLNIVLRDSFAAPEKDIFLALLNW	
ENHAEIMQAVRLPLMSLTELLNVVR	
DAILDAIKVRSESRDMDLNYRGMLII	
MKYGAQVVKGELKSALLDGDTQNY	
FSRHPIDDDCRSGIEIKLGQPSINHVI	
DSRSYSYFIEVSMDELDWVRVIDHS(
WQKLYFPARVCRYIRIVGTHNTVNK ECMFTNKTFTLEKGLIVPMENVATIA	t t
EGVSRSRNALLNGDTKNYDWDSGY	
SGAIVVQLAQPYMIGSIRVLLWDCD	DRSY
541 1891 A 4146 282 778 GTLGYPNGARGQPQDNFFAHQ\VSH	HPPISAC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HAESENFAFWQDMKWKNKFWGKSLEIVPVG
	_					TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG PTPGGQCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR INHLIFRGGAQITFLATTDDSPKAVLGDRLLLT ANVSSENNTPRTSKTTFQLELSVKDAVYTVV SSH
543	1893	A	4153	678	11	TISYPOCLTOMYFLISFANVDTFLLPIMALDH YVAICSALQ*CSIITP/ELCQGLPVLA*AGSSLIS PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF
545	1895	A	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR
546	1896	Α	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029		AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSSPDPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQFPL AASSLFWMFCQPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPAISPPPP AARHRPHPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFFRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSPLGPTSP MP*CFHRPSPPLP/LSSPFPAILRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	201100	ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i	1	ľ	peptide	'	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA
			ł			PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ
		ļ				VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP
						GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG
						LHLPGGRTK
548	1898	Α	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK
j i		ĺ	İ			KIQFHQELLVLFWKLCDFNKVGQPRGALQGD
		1	İ			GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG
			İ			PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR
		ł	1			ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP
						YSIRVPMDIPVFTGTHADLLIV\VFHKIITSGHQ
						RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL
		İ				HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI
]		ļ				IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP
						TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA
			}			PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR
1						MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGOWSPTPEWVLSWKSKLPLO
			}			TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ
						HGTLVGLLPVPHPILIRKYQANSGTAMWFRT
[Ì			YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWORLGVLLSRGKMAVTGWLESLRTAOKT
343	1077	^	4171	976	321	ALLODGRRKVHYLFPDGKEMAEEYDEKTSE
				·		LLVRKWRVKSALGAMGQWQLEVGDPAPLG
						AGNLGPELIKESNANPIFMRKDTKMSFQWRIR
	!		ĺ			NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK
						KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	Α	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL
		-		_		GASAMRRSEVLAEESIVCLQKALNHLREIWE
						LIGIPEDORLORTEVVKKHIKELLDMMIAEEE
						SLKERLIKSISVCQKELNTLCSELHVEPFQEEG
						ETTILQLEKDLRTQVELMRKQKKERKQE\LKL
						LQEQDQELC\EILCMPHYDIDSASVPSLEELNQ
						FRQHVTTLRETKASRREEF/VSSIKRQIILCME
	:					ELDHTPDTSFERDVVCEDEDAFCLSLENIAT\L
j l				[QKLLRQ\LEMQKSQNEAVCEG\LRTQI\RELW
						DRLQIPEEEREAVATIMSGSKAKVRK\ALQ\LE
						VDRLEELEKCKTMKKVIEAIRVELVQYWDQC
ļ l						FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR
						LKNYYEVHKELFEGVQKWEETWRLFLEFER
(KASDPNRFTNRGGNLLKEEKQRAKLQKMLP
						KLEEELKARIELWEQEHSKAFMVNGQKFME
						YVAEQWEMHRLEKERAKQERQLKNKKQTET
·						EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT
}						TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK
]			i	PVAASTCSGKKTPRTGRHGANKENLELNGSI
	:					LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIOS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV
1001	1901	^	4174)	1000	VALLDICFILVAVPESLPEKMRPVSWGAOISW
					l	KQADPFASLKKVGKDSTVLLVCITVCLSYLPE
						AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI
.					l	LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML
						QLAWYGFGSQAWMMWAAGTVAAMSSITFP
					[AISALVSRNAESDQQGVAQGIITGIRGLCNGL
						GPALYGFIFYMFHVELTELGPKLNSNNVPLQ
						GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS
					l	GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS
	·				ŀ	SIWELSSFEEPGNQCTEL
L						

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
цепсе	1	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
552	1902	A	4197	sequence 2	14302	nucleotide insertion
332	1902	^	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR EPAPARRRGTMADGGEGEDEIQFLRTDDEVV
				1		LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
						NSKNVPPDLSICTFVLEOSLSVRALOEMLANT
						VEKSEGQVDVEKWKFMMKTAQGGGHRTLL
1						YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
]						VGLQEDTTGEACWWTIHPASKQRSEGEKVR
1						VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
						QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
						VHARSLWRLETLRVAWSGSHIRWGOPFRLR
						HVTTGKYLSLMEDKNLLLMDKEKADVKSTA
						FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
}						VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
						KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
						TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR
				ł		QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
				ļ		VAGREAGESWKSILNSLYELLAALIRGNRKN
l i					·	CAQFSGSLDWLISRLERLEASSGILEVLHCVL
						VESPEALNIKEGHIKSIISLLDKHGRNHKVLD
1			1			VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
					•	YELMVDHTEPFVTAEATHLRVGWASTEGYSP
						YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG
						CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
						RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
}	}		j		ļ	RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
					·	KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
	i	1		ĭ		KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
				:		ERNYNLOMSLETLKTLLALGCHVGISDEHAE
	J	J	}			DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
						PSQEAMVDKLAENAHNVWARDRIRQGWTY
ļ		l				GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS
1 1	ł	ŀ	l			GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
	İ			İ		GDMRVGWSRPGCQPDQELGSDERAFAFDGF
	l					KAQRWHQGNEHYGRSWQAGDVVGCMVDM
1 1	- [1	ľ		1	NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
1 1	ŀ	- 1	ł	1	}	GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLOEGYEPFAVNTNRDITMWLSKRLPOFLOV
	ŀ	ŀ	}			PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN
		1	1	!		SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
	ŀ	1	1	Ì	l	LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
		- 1	ŀ	į		DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
]]	ļ	j	ł	J	j	RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
[[- [}	1	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHOYDTGFDLDRVRTVTVTLGDEKGKVHE
		-	1		ļ	SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
]]	ļ	j	1		[VVDAASGLLTFIANGKELSTYYQVEPSTKLFP
					ļ	AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
[[- (- 1		ĺ	EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
1		- 1			l	KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
			- 1		ţ	RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
	.	.			ŀ	AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
	1				l	EETKSITLFPDENKKHGLPGIGLSTSLRPRMOF
}	Į	ł	1			SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE
L		1				AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

NO. of pridication of the carbon mucleotide continues of the carbon program of the carbo	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Couling Contragonding Figure Contragonding Figure Contragonding Figure Contragonding Figure Contragonding Figure Contragonding Figure Contragonding Figure Figure Contragonding Figure			nod				
uence 99496 gortspondi ng to first							
ammo acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per							
residue of peptide sequence Poptide sequence V=Tyrosine, X=Unknown, V=Siop pedon, Poptide nucleotide delicin, V=possible nucleotide nucleo	uence			914	ng to first		
peptide sequence muleotide insertion MGIFHNEDLKHILQLIEFSVFKEAATPEEESDT LEKELSVDDAKLQAGGEEAKGGKFKEGLL QMKLPFPVKLQMCLLLQYLCDQVRHEIEAI VAFSDFVAKLQMORGFRYNEWAGALAIMSA ALTARKTKEFRSPQEQMMLLIMFKDIKSECP CPEERIPOFVAKLQMORGFRYNEWAGALAIMSA ALTARKTKEFRSPQEQMMLLIMFKDIKSECP CPEERIPOFVAKLQMORGFRYNEWAGALAIMSA ALTARKTKEFRSPQEQMMLLIMFKDIKSECP CPEERIPOFVAKLQMORGFRYNEWAGAGSVEOPELVR AMFVLLHRQYDGIGGLVRALPKTYTINGVSV BOTNILASLQGISKLISVRMCKEEEKLMIRG LGDIMNNKYFYQHFNLMRALGGHETVMEV AMFVLLHRQYDGIGGLVRALPKTYTINGVSV BOTNILASLGQINSLLSVRMCKEEEKLMIRG LGDIMNNKYFYQHFNLMRALGGHETVMEV WANVLGGISSKEITFFRWANCCRELYCYFCR ISRQNQKAMFPHLSYLLENSSVOLASPAMRG STELDVAAASVMDNIBLALALREPPLEKVVR YLAGCGLQSCQMLVSKGYPDIGWFPVEGER YLDFLRFAVFONGESVEENNAVVVRLLIRPPE CFGPALROEGGNGLLAAMERAIKAEPPSRD GPSINGSSKTLTDTEEEDDTHIMGAAMTFY SALIDLLGRCAPFMHLIHAGKGEARRISILRS LIPLGDLGVGISAFQMPTTAKDGNVYEPDMS AGFCPDHKAAMVLELDRYGGWOPTLHHL EVGFLEPDLRAASALDTAALSATIMALALIRP LCTAVLPLLTRCAPLFAGTEHHASLIDSLHIT VYRLSKGGSLTXAAQROSTEVALLSTGGQLRS MMQHLLRRLVFOVPLINBHAKMPLKLLTHH YERCWKYVCLPGGWONFGAASSEELHLISKK LFWGIPDALSQKKYEGELFSLALPCLSAVAG ALPPD TWISNTYSMMEXGSSMIDSGRIPNPQ PVDTSNTIPEKLEYFINKYAERISHDKWSMMX LANGWYTGGTYSDSSV VQPLMKYFYKLISKE KENTWYTGGTYSDSTATAKTYMKTO LESVKSALRAFIONAGCHALDGTYT ALLPHALSKAKAATVY SQPINN VKPQLLKTHEPTAMETOLOGOG THTNOPPGVAVINGTONAGCHALDGTYN KITCH LCSYMSALRAFALLILDETTLARDL YAFFPLLRFVDYNRAKUKEREOMYVORDMADYNDER HANDLFULGUG							T=Threonine, V=Valine, W=Tryptophan,
sequence Sequence Molenbedt Killin, Liepsynkeaatpeeespit						sequence	
MGPHNEDLKIILQUIEPSVFKEATPEEESDY LEKELSVDDAKLQAGGEEAKGGERKEGLL QMKLPEPVELQMCLLLQVI.CDQVRHRIEAI VAFADDFVAKLQMORGENAGGEEAKGGERKEGEL QMKLPEPVELQMCLLLQVI.CDQVRHRIEAI VAFADDFVAKLQMORGEN YNEWAGALAMSA ALTARKTKEFSSPQEQMMLLNFKDDKSECP CPEERIDQLLDFHEIDLAMFGGIELDEGSLDG NSDLTIRGRILSI.VEKVTVI.KKKQAEKPVES DKKSSTI.QLISETMWRWAGESVEIDPELVR AMFVLLBRQYDGIGGI.VRALPKTYTTNOVSV BOTNALLSI.QQIRSLLSVFMWWAGESKEITHEOPELVR AMFVLLBRQYDGIGGI.VRALPKTYTTNOVSV BOTNALLSI.QQIRSLLSVFMWWAGESKEITHEOLYVREV MYNVLGGESKEITFFRWMANCCHELVYFE LIGDIMANKYFYCHPNILMRALGAHETVMEV MYNVLGGESKEITFFRWMANCCHELVYFE STELDVAAASVMDNHELALAL.REPDLEKYVR VALGCGLQSCQMI.VSKCYPDIGWPYPGGER YLDPILRFAVFCNGESVEERANVVVRLLBRPF CCPPALROEGGNGLLAAMBEAKIA.EDPSRD GPSINSGSSKTLDTEEEEDDTHIMGRAIMTFY SALIDLICRCAFFMHIHAGKGEARIKSIRS LIPICDLVGVISTAPQMFTTAKDGRVVFPOMS AGPEPDERAAAMSLATRAGTEHASLIDSLHT VYRLSKGSSLTKAQRDSEVCLLSICGGURS MAGNILLBRUFVFYFTAKAGTEHASLIDSLHT VYRLSKGCSLTKAQRDSEVCLLSICGGURS MAGNILLBRUFVFYFTAKAGTEHASLIDSLHT VYRLSKGCSLTKAQRDSEVCLLSICGGURS MAGNILLBRUFVFYFTAKAGTEHASLIDSLHT VYRLSKGCSLTKAQRDSEVCLLSICGGURS MAGNILLBRUFVFYFTAKAGTEHASLIDSLHT VYRLSKGCSLTKAQRDSEVCLLSICGGURS MAGNILLBRUFVFYFTAKAGTEHASLIDSLHT VYRLSKGCSLTKAQRDSEVCLLSICGGURS ALTYRITRISQTSTAVAGASEELHLSKK LFWGIPPALSKAKAGTEHASLAVAGA ALPPOYMESN VYSMMEKQSSMDSGGNFNFQ PVDTSMTTIPEKLEFTRIKY ABERIGUKWSMOK LANGWYGEYSDSSKVQPLMKFYKLLSKE KETYR WPFRESLKTMLARTMRTERTREGOSM ALTYRITRISQTSQVSDAAHGVSSRALDMS SNVLSKGLHAALAGAMAENVININWAKKKKM ELESKGGGNHFILVFYDTLTAKERKAKDREKA QDLKFYDEAHQYILLEFGLGGSRGKGEHF PYSGEGEFAKVVLPLIDQYFKNHRLYFLSKA SRPLCSGGMASNIESEGMTSLECKLGVLVRH RISLFGRDATSINCCLHIGGTLDARTYMKTG LESVKSALRAFLDNAAEDLEKTMENLGQGOF THIRNOFKOVTQUINTYTALTPALLSE,EHII GQHQFGEDLILEDVQVSCYRLITSLYALGTSK SYVERGERAAVVENDEAHDYTWAKTG LESVKSALRAFLIDNAAEDLEKKKKAATVV SEEDHLKAARAGROFVALLATHELDGOF THRINGFROUGHAMAKKKKM BLESKGGGNHFILVFYDTLTALEFRALSEHII GCHQFGEDLILEDVQVSCYRLITSLYALGTSK SYVERGERAAVSUPCRAKKHKRGOF VORGENALAFFELILDDEFTLADL YAFYPLLIRRVDYNNEARMELAPPEDEDETKEVL JANVLFHLEGUKSKAVGRRHYCLVEHPQRSK KAWHKILLSKORKRAVACFRMAHYNDER HRAVNLFHQEDELFREDEDFTRENLIEDLA							
LEKELSVIDAKLQĀGGEEAKGGKRFKEGLI. QMKLPPEVALQMCLQYLCDQOVRIEBAM VAFSDDFVAKLQDNQRFRYNEYMQALMMSA ALTARKTKEFRSPQGOMNLLINKFOKSECP CPEERDQLLDFHEDLATHGGIELDEGSLDG NSDLTIRGRILSLIVENVTYLKKQAGKPVES DSKKSSTLQQLISETMYRWQSESVIEDPELVR AMFVLLHRQYDGIGIG VRALPKYTYNOVSV EDTTNILASIQQIRSLLSVRMCKEEKLMIRG LGDIMPNKVFYQHPILMRALGMHEVVMEV MVNVLGGGESKETTPKMVANCCFLCYFCR ISRQNQKAMPPHLSYLLENSSVGLASFAMRG STELDVAAASVMDNIBLALALREPDLEKVVR YLAGCGLGSCMLYSRGYPDIGWRPVEGER YLDFIRFAVFONGSSVEENANVVVKLIRRPF CFGFALRGEGONGLAAMEAKIAEDPSRD GPSNSGSSKTLDTEEEDDTHIMGANIAMTFY SALIDLLGRCAPFMHLHAGKGEARIKSLRS LIPLGBLQVGVISLAFQMPTTAKGGNVYEPDMS AGFCPDHKAAMVLELDRYGGEVQDFLLHLL EVGFLPDLRAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTXAQROFGAASSEELHLISK LFWGIPDALSQKKYFQELFKLALPCLSAVAG ALPPDVMSSTYVSMMEKQSSMDSSGNPNPQ PVDTSNTTIPEKLEYFINKYAEHSIDKWSMKK LANGWYGGFYSSSKVQDLMKYKYLLISKEE KEIYRWPIKESLKTMLARTHRERTREGDSM ALYNRTRRSOTSGVYDAAHGVSENFIND NVTLSRDLHAMAEMMAENYHNIWAKKKM BLESKGGONHPLLVPYDTLTARKAKKREA QDILKFLONGVAVSRGFKDLELDTSEKREA YSLQQLIKTVODEAHQVILEPGGSRGKGEHF PYEGEKFFAKVUPLIDQYKKNIRLYFLSAA SRPLCSGGHASNEKEKMYLEPGIDARTWKKTC LESVSALARALDNAARMAEMMAENYHNIWAKKKM BLESKGGONHPLLVPYDTLTARKKAKREEA QDILKFLONGVAVSRGFKDLELDTSEKREA YSLQQLIKTVODEAHQVILEPGGSRSKGCGHF PYEGEKFFAKVUPLIDQYKKNIRLYFLSAA SRPLCSGGHASNEKEKMYSLFCKLGVLVAH RISLEGDDATSIVNCLHILGGTLDARTVMKTG LESVSALARALDNAARASIPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETITARD KNYTTSRDLHAMAEMMAENYHNIWAKKKKM SIVYERQSSALGGCLAAFAGAPVAFLETHLD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD KHNYSTYNTKSSERAALSJPTNWEDVCPHIP SLEKLMEENVELAEGIRVTOMPITYMETURD ARWONGREDLILEDLAKKKAATVY SQPIBKVROPLIKTHIPLEGIRCACAPOCELIA LANRESLKOTEGEDURDINHOLGCLOD ARWONGRAVSKOPRINGVLOFMENN MSELITITISKSKNORRHYCLVEHPORSK KAWMHK					Sequence		
OMKLPEPVKLOMCLLLOYLCDCOVRINGS VAFSDDFVAKLOMORRFNYEWOMALNMSA ALTARKTKERSPPOGOIMMLINKDDMSEC; CPERIBOQLLOFHEIDMTHICGIELDEGOSLDG NSDLTIRGRILSLVEKYTYLKKQAEKPVES DSKKSSTLOQLISETMYRWAQESEKUMER DSKKSSTLOQLISETMYRWAQESEKUMER LSDMYLLHROYDGIGGLVRALPKTYTHOLYSY EDTINLASLOGIRSLSVEMGREEKUMER LDDIMNNKVFYQIPPNLMRALGMHETVMEV WAVNVLGGGESKETTPEMVANACCFRLCYFCR ISRONOKAMPDHLSYLLENSSVGLASPAMBG STELDVAAASVMDNINELALARFPILEKVYR YLAGGGLOSOMLVSKGYPDIGWAPPVEGER TYLDELRRAVPCHORDNINELALARFPILEKVYR YLAGGGLOSOMLVSKGYPDIGWAPPVEGER YLDELRRAVPCHORSVEYENANVYRLLIRRPP CFGPALRGEGYMOLLAAMFEAIKLAEDPFOM GFSTENSOSSKLTDTEEEDDITHINGAMAPTFY SALIDLIGRAPEWHILHAGKGEARRISILSE LPLGIGVSISAFQMFTIAKOGNVVEPDMS AGFCPDIKAAMVLFLDRYYGIGVOPPLHILL EVGTLPDLAAASSLDTAALSATDMALALINFY LCTAYPLERAAASLDTAALSATDMALALINFY LCTAYPLERAAASLDTAALSATDMALALINFY VRILSKOCSLTKAQROBEVCLLSICGQLPR MMQHLIRRLYPDYPLLINBHAKMPLRLITNH YSRKGKSTLTAQROBEVCLLSICGQLPR MMQHLIRRLYPDYPLLINBHAKMPLRLITNH YSRKWYYCLPGGWGMFOAASEELHILSKK LPWGIPDALSQKKYEGELRKLALPCLISAVAG ALPPDYMESSNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPBELE FYNNIV ABBINSWSMMK LANGWYGEISTSSNYVSWDAAHGYSPRAIDMS NITISSDLISAMGMAENTHINAKKKKM BLESKGGGNIPLLYPYDTLTAKERAKOREK KEITYMPRESLICHAARTHREGDSM ALYNGRTRISGTSOVSVDAAHGYSPRAIDMS NYLSSDLIAMMERMANNIVAKKKKM BLESKGGGNIPLLYPYDTLTAKERAKOREKA GDILKRYWRIESLICHALRYTHREGOSM ALYNGRTRISGTSOVSVDAAHGYSPRAIDMS NYLSSDLIAMMERMAENTHINAKKKKM BLESKGGGNIPLLPYDTLTAKERAKOREKA GDILKRYMFRESLICHALRYTHKTOL BERNARDHANAKCHURTH RISLGFONATSINVASRGKFOLELDITSPIEKRRA YSFLQULRYVDRAHGYTLISKERA YSFLQULRYVDRAHGYTLISKERA SRPLCSGGHASNKEKEMVYSLFCKLGVLVARH RISLGFONATSINVASRGKFOLELDITSPIEKRRA YSFLQULRYVDRAHGYTLISKERH GOHGGEDLIEDWQVSCYRLITSLYALGTSS SIVVERGRALGECLAAFAGAFPVAHETHLD KHNYSTYNTKSSRERAALSLFINNEDVCPHIP BLEKLMABAGNARHINNUTIVALLIPHE HERMANLFLQNYELSSWIFTEEHHINGKLKKAATVY SQPINNVKYRLLSKRIFKREGDNYMKRLAVYF SQPINNVKYRLLSKRIFKREGDNYMKRLAVYF SQPINNVKYRLLSKRIFKREGDNYMKRLAVYF SQPINNVKYRLLSKRIFKREGDNYNVONENN MSFLITDTKSKMSKAAVSDGEKKMKRKGD PANGTSLIVALLFSRTALT UANNFLLKOFFOULLERHERE LAKHRISKLISKREGMYTOWNENNUTHEN HYNDYTILDRYNKULLSKREDHETTLARDL JANNFLLSKOFERHYCLUE							
VAFSDDEVAKLQDNQRFKYNEVMQALMMS ALTARKTKERSPGEORMLINKDDKSEC CPERRDQLLDFHEDLMTHICGIELDEDGSLDG SIDGUTRIGRILIS VER VYTIKKÇAEKPVES DSKXSSTLQQLISETMYR WAQESVIEDPELW; AMFVYLHRQYDGIGGURALPKTYTHOVSY EDTINLLASLQGISLLSVRMGKEEKLMIRG LDDIMNKLYPQUPINLMBALGMHETYMEV MYNVLGGGESKETTPKWANCCRFLCYNC SISRQNQKAMFDHISYLLENSSVGLASPAMEG STPLDVAAASVMDNNELALALREPDLEKVYR YLAGCGLQSQMLVSKGYPDIGWRPVEGER YLDELRAYPCRGGSVEENANVYNLLIRRPE CFGPARGEGGNGLLAAMEEAKHAEDPSD GISSNSSSKILDTEEEDDTHIMGAMITTY SALIDLLGRCAPEMHLIHAGKGEARIRKSIRS LIPLGDLVGVISLAGYMFTIAKOGNAMITTY SALIDLLGRCAPEMHLIHAGKGEARIRKSIRS LIPLGDLVGVISLAGYMFTIAKOGNAMITTY VARISKGSLTKARQDSISVCLISICGQLRPS MMGHLIRRLYDVPLLNEHAKMPLKLLTHT VYRLSKGCSLTKARQDSISVCLISICGQLRPS MMGHLIRRLYDVPLLNEHAKMPLKLLTHN YERCKYYCLPGWGNFGAASEGHING LPWGIPDALSQKKYEQELEKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFING PVDTSINTIPEKLEYFINKYAEHSHDKWSMDK LANGWYTGGIYGDSKVQPLMFYKLLSKKE KEIYRWPIKESLKTHMLARTMRTERTREGDSM ALYNRTRRISGTSSSKVQPLMFYKLLSKKE KEIYRWPIKESLKTHMLARTMRTERTREGDSM ALYNRTRRISGTSSVSVODAHGYSFRAIDMS NYTLSDDLHAMAEMMAENYHNIWAKKKRNEKA QDILKFLQINGYASKGCHEP PYEGEIKFFAKVYLPLDQYFKHHLISKE LESKGGANFIPLLYPYDTLTAKERMILKQDEA ASPICLGGGGRAPHILAGYLDFORDER STYPEGREFAKVYLPLDQYFKHHLISSFEHI GOHQFGEDLLEDVYNOSTRUTHALSSFEHI LESKGGGRAPHLYPYDTLTAKERMILKQOOP THIRNQPKGVTQINNYTTVALIFMSLSFEHI GOHQFGEDLLEDVYNOSTRUTHALSSFEHI LESKGGRAPHLYNFULLYBLSSFEHI LESKGGRAPHLYNFULLHERSLGCOP THIRNQPKGVTQINNYTTVALIFMSLSFEHI GOHQFGEDLLEDVYNOSTRUTHALGTLAARTWRICHTLAGD LESKKSARAFLDNAADELEKTMERLKQOOP THIRNQPKGVTQINNYTTVALIFMSLSFEHI GOHQFGEDLLEDVYNOSTRUTHALGTLAARTWRICHTLAGO THIRNQPKGVTQINNYTTVALIFMSLSFEHI GOHQFGEDLLEDVINNYTNAKKKENNENELKQOOP THIRNQPKGVTQINNYTTVALIFMSLSFEHI GOHQFGEDLLEDVINNENKENENENEVYONENN MSFLITDTKSKMSKAAVSQUEKKKMKKGO RYSNGTSLIVAALKRENNERAESORYVONENN MSFLITDTKSKMSKAAVSQUEKKMKKKGO RYSNGTSLIVAALKRENNERAESORYVONENN MSFLITDTKSKMSKAAVSQUEKKMKKKGO RYSNGTSLIVAALKRENNERAESORYVONENN MSFLITDTKSKMSKAAVSQUEKKMKKKOOP ARWOMALYKLIPHORIDDTSDEKFRALLOGLUED ARWOMALYKLIPHORIDDTSDEKSTERALLTORLOGHED ALKOMALYKLILBORDAADLEKTMATOROOPELLED ALKOMALYKLILBORDAADLEKTMEREDENEVYONENN MSFLIT							
CPERIRDQLLDFHED.MTHICGIEL DEDGSLDG NSDTIRGRILIS. VER VYTYLKKOAREPVES DSKKSSTLOQLISETAVR WAQESVIEDPELY AMFYLLHRQYDGIGGL VRALPKTYTHOK VSV EDTINILASLOQIRSLS VRMGKEEEKLMIRG LODIMNIKVPYQHPNLMRALOMHETVMEV WYVLGGGESKEITPFEMVANCERCHCYFCR ISRQNQKAMFDHLSYLLENSSVGLASPAMRG STEPLDVAAASVMDNNELALALREPDLEKVVR YLAGGGLQCQMLVSKGYPDIGWNPVEGER YLDELRRAYPCRGSSVEENANVVALURRPPE CFGRALRGEGGNGLLAMEEAIKIAEDPSRD GFSRNGSSKTLDTEEEDDTHIMGAIMTEY SALIDLLGRCAPEMHLHIAGKGEAIRISLIRS LIPLGDLVGVISIAGRAPMTHAKDGHVALLEN LEGGLVGVISIAGRAMVILDRYVGIEVQDFLLHL EVGFLPDLRAAAS LDTAALSATDMALALINRY LCTAVLPLLTRCAPLPAGTEHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRFS MMQHLLRILVEDVPLINBHAKMPLKLLTNH YERCWKYYCLPGGWGNFGAASEELHLISK LPWGTDALSQKKVEQELKKALJCLSAVAG ALPPDYMESSYVSMMEKQSSMDSEGNFNPQ PODTSYNTIPEKLEVFINNYABEHSEMSWMDK LANGWYGGEYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKIMLARTMRTERGOSM ALYNTRISIGTSQVSVDAAHGYSPRAIDMS NYTLSRDLIAMABEMMAENYHTINVAKKKM BLESKGGGNPPLLVPYDTLTAKEKAKDREKE AUGHTENDATASHANGAKKM BLESKGGGNPPLLVPYDTLTAKEKAKDREKE ODILKFLONDYAVSGRFALELDTFSERRPA YSFLQQLIRVYDEAHQYLLEPDGGSRGKGEHF PYEGGEKFFAKVVLPLDOYKFNHFYLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVWH RISLFONDATSIVNCLHLGQTLDARTVMKTO LESVKSALRAFLDNAAFDLEKTMENKOGOF THITNIQKSVTSJENTALLFONDARTVMCHO KNINTSYNTINTSKERERAALSLETNUNGEVLSAA SRPLCSGGGALSNKEKEMVTSLFCKLGVLVWFO SEEDHLKABARDAMAEDLEKTMENKOGOF THITNIQKSVTSJENTALLFONDARTVMCHO LESVKSALRAFLDNAAEDLEKTMENKOGOF THITNIQKKAKKAATVV SEEDHLKABARGDMSKEALLILDEFTILARDL KNINTSYNTNINTSKSRERAALSLETNUNDEVCPNIP SLEKLMEEVELAESGIRYTQMPIVMEVULPM LCSYMSWWEHGEPENPFRABELFR MYAEVITYWSKSHNIKEEDONFVONDENN MSFLITDTTKSKMSKAAAVSOQERKKAATVV SEEDHLKABARGDMSKEALLILDEFTILARDL VAFFYLIRFVDVNAKKWROD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNFSKLKERMYSKEEDONFVONDENN MSFLITDTTKSKMSKAAAVSOQERKKKARVOD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNFSKLKEEMVTSKEENKEERGENFFVONDENN MSFLITDTTKSKMSKEEMALTNIPRICHALGOLICAPGDQELIA LAKNFSKLKTHIPLIPMELLELLLFSRTALT DIANVIFILBCKSKRAVASCPERKEKAATVU			,				
NSDLTRIĞRLISLVEKYTLIK(KQAEKPVE) BOKKSSTLQQLISETMYRWQESVEDPELVR AMFVLLIR(Y)TGIGGLVRALPKTYTINGVSVE EDTINLASLGQIRSLLSVEMGEKELMIRG LDIMINIKVFYQUPNLMBLALGMHETVMEV MVNVLGGESKEITFYKMVANCCRILCYFCR ISRQNQKAMPDIHSYLLENSSVGLASPAMRG STPLDVAAASVMDNNELALALREPDLEKVVY VLAGCGLQSCQMLVKSKYPTDIGWNPVEGER YLDELRFAVFCNGESVEENANVVVRLLIRFSP CFGPALGEGGGLAMEEALKIAEDPSRD GFSINSGSSKTLDTEEEEDDTHIMGNAIMTTY SALIDLLGRCAPPAHILHAGKGEARRISLRS LJPLGDLVGVISIAFQMFTIAKDGNVVEPDMS AGFCPDIKAAMVLFLDRVYGEVPCLHLLL EVGFLPDLRAAASLDTAALSATDMALALRIPSRD UCTAVIPLLTRCAPLAFDRYGHERVSLIDLLHL VYRLSKGCSLTKAQRDSIEVCLLSICGQLRFS MMGHLIRRLVFDVPLINBHAKMPLKLLTNH YFRCWKYYCLPGGWGNFGAASEELHHISK LFWGFDALSQKKVQELFKALALCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFTPD PVDTISTNITJEKLE VFINNYABHSHDWSMDK LANGWYGGIYSDSSK VQPLMKPYKLLISEKS KEIYRWFKESLLTMLARTIMTERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILLVFTOTTTAKERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILVFTOTTTAKERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILVFTOTTTAKERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILVFTOTTTAKERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILVFTOTTAKERGISM ALYNTTRISQTSQVSVDAAHGYSFRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGHPILVFTOTTAKERGISM ASPLCSGGGHSNIKGHPINTTVALLFMLSGLFCH PYEGEKFFAKVVLPLIDQVFKNHRLYPLSA SISVERQRSALGELCLAPAGAFPVAFLETHLD HINNOPKGYTQIINTTVALLFMLSGGF THITNOPKGYTQIINTTVALLFMLSSLFEHI GOHGGEDLLIEDVQVSVCRILITSALTALGTSK SISVERQRSALGECLARFAGAFPVAFLETHLD KINTYSINTKSSRERAALSLTPTNEDUCPNIP SLEKLMEEVELAESGIRYTOMPIVVMEVLIP LCSWMSWWEHGEFRONPERABERIF NVAEVFULIRVVVNSKSINFKREQNFVVQNENN MSFLITDTKSKMSKAAVSDGERKKMKRKOD RYSMGTSLLVAALKRATVUSSEEDHLARAARGDMSEAFLILDEFTTLARDL VAFFYLLIRFVDTNIKAKSKENTUSPECTUSPEL DIANVLFHLEGKSKRAVACFREAELIF NVAEVFULREVDVINARKULLEDLAR KARPSLKDTILDVKKMSKRAAVLOFDREAELIF RWAMLYKLJUGNYKKSWETEEHYFFEDKLLEDLA KRAPSLKDTIDTKSKMSKAAVSDGERKKMKRKOD RYSMGTSLLVAALKSKRAAVLOFTREAELIF LAKRRSLKDTIDTKSKMSKAAVSDFREEDIF HRAVNIFLLGOYEKSWIETEEHYFEDKL							· · · · · · · · · · · · · · · · · · ·
DSKKSSTLQQLISETMYR AQESVIEDPLVR AMPVLHRQYDGIGGLVRALPKTYTINGVSV EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG LGDIMNINVYOHPMLMRALGMHETYMEV MVNVLGGGESKEITFFKMVANCCRFLCYFCR ISRQNQKAMPHLSYLLENSSVGLASPANRG STPLDVAAASVMDNNELALALEPDLEKVYR YLAGGGLQSCQMLVSKGYPDIGWNPVEGER YLDFLRFAVFONGESVEENANVVRLLIRRPS CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD GFSPNSGSKTLTDTEEEDDTIHMGNAIMTFY SALIDLLGRCAPFMHLHAGKGEARIRISILRS LIPLGDLVGVISIAFOMPTIAKANVVVPLINRPS AGFCPDHKAAMVLFLDRVYGEVOPCHLHL EVGFLPDLRAAASLDTAALSATDMALALNIY LCTAVLPLLTRCAPLFAGTEHHASLIDSLHT VYRLSKGCSLTKAASIDTAALSATDMALALNIY LCTAVLPLLTRCAPLFAGTEHHASLIDSLHT VYRLSKGCSLTKAAGNDISEVCLSICGOLRPS MMQHLLRILVFDVPLINEHAKMPLKLLTDH YERCWKYYCLPGGWGNFGAASEEELHLSK LIPWGFFDALSQKKYFQELFKLALPCLSAVAG ALPPDYMESKYVSMMEKQSSMDSGGNFRPQ PVDTSNTIFPEKLFYNKYAEHSHDKWSMMS LANGWYGEIYSDSSKVOPLMKPYKLLSEKE KEINTWPRESLKTMLARTIMETERTEEGDSM ALYNSTRRISQTSQVSVDAAHGYSPRAIDMS NYTLSRDLHAMAEMMAENHINVAKKKM ELESKGGGNIPHLLVPYDTLTAKEKAKDREKA QDILKFLQNGVAVSRGFKDLETPSIEKRFA VSFILQQLIRYVDEAHQYILLFDEGGSRGKGEHF PYEGGIKFFAKVYLPLIDQYFRHIYMAKKKM ELESKGGGNIPHLLVPYDTLTAKEKAKDREKA QDILKFLQNGVAVSRGFKDLETPSIEKRFA VSFILQQLIRYVDEAHQYILLFPGGSRGKGEHF PYEGGIKFFAKVYLPLIDQYFRHIYMSLYSLFHH QCHQFGEDLLIEDVQVSYCRIFTLYPSIEKFRA SPPLCSGGHASNEKERMYTSLFCKLGVLVRH RISLFGNDATSIVCLHHIGQITTSLYALGTSK SITYBERGRSALGELAAFAGAFPVAFLETHLD KINNISYINTIKSSRERAALSLIPTINVEDVCIPHD SLEKLMEEIVELAESGIRYTQMFIYMEVLIPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMMTLLGRILKHIYNINLGIDEGAWMKRLAVF SQFINKVKPQLLKTHIFLIRMSLKKKKAATVV SEEDHIKAEAGRGMSEAELLLLDEFTTLARDL YAFFYBLIRFTVDNISK KARAD LAKNFSLKORGENFOVONDENN MSFLITDTKSNKKAKAAYDOPROKKKKKAGD RYSMGTSLUVAALKRLIPGIAICAPDQELIA LAKNRFSLKORGENFOVONDENN MSFLITDTKSNKKKAAYDOPROKKKKKAGD RYSMGTSLUVAALKRLIPGIAICAPDQELIA LAKNRFSLKORGRAVAVGCREPVCULEDD AIRWGMALYKOLPRIRDSTIPCKTERUL DIANVLFHEQKSKRVGRRHYCLVEHPORSK KAVWHKLLSKORKRAVVAGRRHYCLVEHPORSK KAVWHKLLSKORKRAVVAGRRHYCLVEHPORSK KAVWHKLLSKORKRAVORRHYCLUELDLA KPGAEPPEEDEGTKKVDFLHOPLLEDLA							l
AMFVLHRÖYDGIGGLYALİPKTYINGVSV EDTINLIASLGQIRSLLSVRMÖKEEEKI.MIRG LGDIMINNIX YFYOHPNI MRALGMHETYMEV MYNVLGGGSKETIFFKMYANCCEPLCYPCR ISRONOKAMFDHLSYLLENSSVGLASPAMRG STPLDVAAASVMDNNELALALREPDLEKVVR YLAGGCLOSCOMU.NSKGYPDIGWNFVEGER YLDFLRFAVFCNGESVEENANVVVLLIRFE CFGPALRGEGGNGLAAMEEARIKAEDPSRD GFSPNSGSSKTLDTEEEEDDTIHMGMAIMITY SALIDLLGRCAPMHIHAGGARIRSLRS LIPLGDLVGVISIAFOMPTIAKDGNVVPEDMS AGFCPHKAAMVLFLDRYVGEVOPFLHIL EVOFIPDLRAAASLDTAALSATDMALALINY LCTAVLPLLTRCAPLFAGTEHHASLIDSLHT VYRLSKGCSLTKAQRDSIEVCLISICGQLRPS MMGHLRRLVFDVPLINEHAKMPLKLLTINH YERCWKYYCLOPGGWNFOAASEELHLSRK LFWGFPDALSQKKYEGGLFKLALPCLSAVAG ALPPDYMESSYVSMMEKOSSMDSEGNFNPQ PUDTSNTITIPEKLEYFINKYAEHSHDKWSMOK LANGWYGGIFYSDSKVOPLAKHYKLLSKE KEINRWPKESLKTMLARTMSTERTEGDSM ALYNTRISIOTSONSVOYDAAHSYPRALDMS NYTLSRDLHAMAEMMAENYHHIWAKKKM ELESKGGGNIPPLLVPYDTLTAKEKAKDEKKA SUPLSKGGGHIPPLLVPYDTLTAKEKAKDEKKA QDLKFLQNOYAWSRGFKDLELDTPSIEKRRA YSFLQQLIRYVDEAHQYLLEFFGGSRGCHIP PYEGGIKFFAKVYLPLLDQYKRHIRLYFLSAA SRPLCSGGHASNKEEEMVTSLFCKLGULVRH RISLFONDATSINCHLINGGURATVMETG LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVMETG LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVMETG LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVESTIC LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVESTIC LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVESTIC LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATVESTIC LESWSAALRAFLDNAAEDLEKTMENLKQGOF THITRIOPKGVTVINCHLINGGURATALETHLD KINNISTINTKSSEERAALSLFTNVEDVCPNIP SLEKLMEEIVELAESGIRVTQNMITWEVILPM LCSYMSRWWHENPENNEERAEMCCTALINGE HMYTLLGRILKINTNILGIDEGAWMKRLAVP SQPIDINKVKPQLLKTHFLILMERLKKKAATVV SEEDHLKAEARGORMSEAELLLLDEFTTLARDL YAFYPLLIRFYDYNRAKWIKERGMCTALANDL ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWRESKENTEDEVRURINGHLICKLEDD ALKWPELLTDLICKRRACKRACHTCULLEFRORSK KAVWHKLLSKORKRAVVACREMELLLLA	l i					"	
EDTINLLASI.GQIRSILSYMGKEEEKI.MIRG LGDIMNNNY FYOHPIM.MRAI.GMHETYMEY MYNVLGGGESKEITFYKMVANCCRFI.CYFCR ISRQNQKAMFDHLSYILENSSVCIAANRG STPLDVAAASVMDNNELALALREPDLEK.VYR YLAGGGLQSCQMLVSKGYPDIGWNFVEGER YLDFLRFAVF.ONGESVEENANVVRLIJRRYF CFGPALRGEGGNGI.LAAMEEAIKIAEDPSRD GFSPNSGSSKT.DTFEEEDDTINGAMITY SALIDLI.GRCAPPMH.IHAGKGEARIRSILRS LIPL.GDLVGVISIAFOMPTIAKOMVEPDMS AGFCPDHKAAMVI.FLDRVYGIEVQDFLI.HL EVOFLPDLRAAASLDTAALSATDMALALINFY LCTAVLFLITR.CAPLFAGTEHALDISDLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHL.RRL.VFDVFLINEHALDISLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMGHL.RRL.VFDVFLINEHALDISLLHT YFRCWKYYCLPGGWOFGAASEEELHLSRK LFWGFDALSOKKYFGCEFLKALDSSLLHT YFRCWKYYCLPGGWOFGAASEEELHLSRK LFWGFDALSOKKYFGCEFLKALDSSLLHT VRIJSKGCSLTKAQRDSIEVCLLSICGGLRPS MAYDYMESN.VSNMEKQSSMDSEGMFNQ PVDTSNITIPBEK LFYFINKY ABLOFLSAVAG ALPPDYMESN.YSNMEKQSSMDSEGMFNQ PVDTSNITIPBEK LFYFINKY ABLOFLSAVAG ALPPDYMESN.YSNMEKQSSMDSEGMFNQ PVDTSNITIPBEK LFYFINKY ABLOFLSAVAG ALPPDYMESN.YSNMEKQSSMDSEGMFNQ PVDTSNITIPBEK LFYFINKY ABLOFLSAVAG ALYNTRIRSIGTSQVSVDAAHGYSPRAIDMS NYTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGONIPLLVPYDTLAKFKARDFEKA QDILKFLQNGYAVSRGFKDLELDTPSIERRPA YSFILQLIRYVDEAHQYLEKFRADERRA SPPLCSGGHASNEKEEMVTSPLGGSRGKGEHF PYEOEIKFFAKVULPLIDQYFKNHRLYFLSAA SEPLCSGGHASNEKEEMVTSPLGGSRGKGEHF PYEOEIKFFAKVULPLIDQYFKNHRLYFLSAA SEPLCSGGHASNEKEEMVTSPLGGSRGKGGHF PYEOEIKFFAKVULPLIDQYFKNHRLYFLSAA SEPLCSGGHASNEKEEMVTSPLGGGSRGKGGHF PYEOEIKFFAKVULPLIDQYFKNHRLYFLSAA SEPLCSGGHASNEKEEMVTSPLAGOOF THITNRQFKGYTQINYTTVALIPMLSSLFFHI GOHQFGEBLILEDVOXSYRIT.TSLYALGTSK SIYVERGRSALGECLAAFAGGATALETHLD KHNYSYYNTKSSRERAALSLFTNVEDVCPNIP SLEKLMEEIVELAESGRYTOMYMEVULPM LCSYMSRWWEHOPENNPERAEMCCTALNSE HMMTLLGNILKIIVNNLGIDEGAWMKRLAVF SQPINKVRQLLKTHFLPLMEKLKKAAATVV SEEDHLKABARGAMSKAAVSOGERKMKRGOOF THITNRQFKGYTUWSKSINNFREEQNFVVONENN MSFLITDTKSSKNKAAVSOGERKMKRGOOF RYSMGTSLIVAALKRLLPGILNICAPGDQELIA LAKNFSLKDTEDEVBUIRSNICKACATVU SEEDHLKABARGARVALAKRLAUPULLLEBLA KYGAGEPPEEDEETKEDENLIEDLA KPGABPTEEDEGTKKVDFLHQLLLLFSRTALT HRAVNLFILQFYERSWIETEHYYEDKLIEDLA KPGABPPEEDEGTKKVDFLHQLLLLFSRTALT							1
LGDIMINKYF\GHPILMRALGMHETVMEY MVNVLGGESKETIPFKMYANCERLCYFCR ISRONQKAMFDH.SYLLENSSVGLASPAMBG STPLDVAAASVMDNVELALALREPDLEK VVR YLAGGGLQSCQMLVSKGYPDIGWNFVEGER YLDFLRFAVFCNGESVERANVVVRLIERRFE CFGPALRGEGGNGLLAAMEAINTPY SALIDLLGRCAPMELHHACKGEARIRSILRS LIPLGDLVGVISIAFQMPTIAKDGNVVEPOBMS AGFCPHKAAMVLFLDRVYGEVQDFLHLL EVGFLPDLRAAASLDTAALSATDMALALNSY LCTAYJPLITRCAPLFAGTEHALISDLHT VYRLSKGSLTKAQRDSIEVCLLSICGQLRPS MMQHLRRLVFDVPLNEHAKMPLKLLTNH YSRCWKYYCLPGGWGNFGAASEELHLSRK LPWGFDALSQKKYEQGLFKLALPCLSAVAG ALPPDYMESSYVSMMEKQSSMDSEGNFNPQ PVDTSNTITPEKLEYFINKYAEHSHDKWSMDK LANGWYGEIYSDSSKVQPLAKHJSKEK KEINRWPKESLKTMLARTMETERTEEGDSM ALYNTRISIGNGVSVDAAHGYSPRAIDMS NYTLSRDLHAMAEMMAEMYHNIWAKKKM ELESKGGGNIPFLVPVDTLAKEKARDREKA QDILKFLQNGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEPDGGSRGKGEHF PYEGGIKFFAKVYLPLIDQYKRHIKLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFONDATSINVCHLIGGTLATVMKTG LESWSALRAFIDNAAEDLEKTMENLKGGOF THITRIOPRGVTINVCHLIGGTLATVMKTG LESWSALRAFIDNAAEDLEKTMENLKGGOF THITRIOPRGVTINVCHLIGGTLATVMKTG LESWSALRAFIDNAAEDLEKTMENLKGGOF THITRIOPRGVTINVCHLIGGTLATVMKTG LESWSALRAFIDNAAEDLEKTMENLKGGOF THITRIOPRGVTINVCHLIGGTLATVMKTG LESWSALRAFIDNAAEDLEKTMENLKGGOF THITRIOPRGVTINVTTVCHLIJSSLFFHI GOHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERGRSALGGCLAAFAGRAVLETHDL KHNISTYNTKSSERAALSI-PTNVEDVCPNIP SLEKLMEEIVELAESGIR YTQMPITVMEVILPM LCSYMSR WWHEHPENNPERAEMCCTALANSE HMYTLLGRILKHTNNLGIDEGAWMKRLAAVP SQEPILNKVKPQLLKTHEILLMEKLKKKAATVV SEEDHLKAEARGMSSAELLLDEFTTLARDL YAFYPLIRFYDYNRAKWILKEPNIEGENN MSFLITDTKSKNKKAANSDOGERKMKKRGD RYSMGTSLIVAALKRLLPGLINCAPDGDELIA LKKNPSI-KDTEDEVRURINSHIHLQCKLEDD AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSRKVORRHYCLCVERPORSK KAVWHKLLSKORKRAVVORGREMENLEDLA KPGARPPEEDEETKRVDPLHOPERSE							-
ISRONOKAMPDHLSYLLENSVOLASPANRG STPLDVAASVMDNNELALREPDLEKVVR YLAGCCLQSCQMLVSKGYPDIGWNPVEGER YLDFLFFAVFCNGESVEENANVVYRLIRREF CFOPALROGGORGILAAMEAIKIAEDPSRD GPSPNSGSSKTLDTEEEEDDTIHMGNAMTFY SALIDLLGRCAPEMHLHAGKGEARIRSILRS LIPLGDLVGVISIAPQMPTIAKDGNVYEPDMS AGFCPDHKAAMVLFLDRVYGIEVQDFLHLL EVGFLPDLRAAASLDTAAASTDMALALNRY LCTAVLPLLTRCAPLAGTEHHASLIDSLHT VYRLSKGCSLTKAQRDSIECLISICGQRPS MMQHLLRLVFDVPLLNEHAKMPLKLLTHH YERCMKYVCLPGGWORGASEEGLHLSRK LFWGIFDALSQKKVEQELFKLALPCLSAVAG ALPPDYMESSTVYSMMEKGSSMDSEGNFPNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIVSDSSKVQPLMKFYKLLSEKE KEIYKWPIKESLKTMLARTHERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSRAIDMS NVYLSRDLHAMAEMMARNYINIWAKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGVAYSRGFELDLTPSIEKRFA QVSLKAVALVARTRAKSGTSQVSVDAHGYSRLISKE ELSKGGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGVAYSRGFELDLTPSIEKRFA YSFLQQLIRYVDEAHQYLLEFDGGSRGKGEHF PYEQEBKFFAKVVLPLIDQYFKNIRLYFISAA SRPLCSGGHASNKEKEMYNSTLFCKLGVLVRH RISLFGNDATSIVNCLHLQGTLDARTVMKTG LESKYSALRAFLDNAAEDLEDTPSIEKRFA YSFLQQLIRYVDEAHQYLLFTSAA SRPLCSGGHASNKEKEMYLTSLAYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLFINVEGOG THITRNOPKGVTQINYTTVALLPMLSSLFBHI GQHQFGGDLILEDVQVSKYTLSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLFINVEDVCPNIP SLEKLMEEDVELAESGIRYTSLAYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLFINVEDVCPNIP SLEKLMEEDVELAESGIRYTSLAYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLFINVEDVCPNIP SLEKLMEEDVELAESGIRYNTENFREREMCCTALNSE HMMTLLGNILKITYNLGDEGAWMKRLAVF SQPINKVKPQLLKTHFLPLMEKLKKAATVV SEEDHLKAERGROMSEABLILDEFTTLARDL YAFYPLLIRFVDYNRAK WLKEPNPEAEHCFCALINGE HMYBLGRIKKKAATVV SEEDHLKAERGROMSEABLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEHCFCALINGE HMYBLGTIDTISKMISKNIHLOGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLIGENICAPGORGLIA LAKNRFSLKDTEEDHYPEINSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLIGENICAPGORGKEL KOGABPPEEGEGTKKVDPLYCLLFERPRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HAVNLFLQGYEKSWIETEEHYPEEDKLIEDA KROGSPEPEGGTKNTDPLHQLILLFSRTALT	1						
STPLÖVÄNASYMDNNELALALREPLEKVYR YLAGCGLOSCOML VSKOPIGWNPVEGER YLDFLRPA VFCNGESVEENANVVVRLLIRRPE CFOPALROGGORGLAAMELAIKADPSRD GFSPNSGSSKILDTEEEEDDTHIMGNAMTTY SALIDLLGRCAPEMHLIHAGKGEARRISLRS LIPLGDLVGVISIAFQMTHAKDGNVEPDMS AGFCPDHKAAMVLFLDRVYGIEVOPDLHLL EVOFLPDLRAASLDTAATMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKILTNH YERCWKYYCLPGGWGNFGAASEEFLHLSRK LIPWGIFDALSQKKVEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEONFNPQ PUDTSNITIPEKLEYFIRNYAEHSHDKWSMOK LANGWYGEIVSDSSKVQPLMKPYKLLSEKE KEIYRWFIKESLKTMLARTHETRTEGOSM ALVNRTRRISQTSQVSVDAAHGYSPRAIDMS NYTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGONHPLLPYPTITAKEKAKDREKA QDILKFLQINGYAVSROFKDLELDTPISEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKIFFAKVVLPLIDQYFKNHRILFLSAA SRPLCSGGHASNKEKEMYSLFFLGGOF THITRNQPKGVTQINYTTVALLPMLSSLFHI GGHQFGEDLLLEDVQVSCYRILTSLYALGTSK STYVERQRSALGECLAAFAPVAFLETHILD KHNIYSIYNTKSSRERAALSLPINVEGOP THITRNQPKGVTQINYTTVALLPMLSSLFHI GGHQFGEDLLLEDVQVSCYRILTSLYALGTSK STYVERQRSALGECLAFPAVVAFTTHLD KHNIYSIYNTKSSRERAALSLPINVEDVOPNIP SLEKLMEEPVELAESGIRYTOMPIVMEVILPM LCSYMSRWWEHGFENPERAEMCCTALNSE HMYTLLSGNLKINYNAKTG SEEDHLKAEARGDMSSEALLILDEFTILARDL YAFYPLLIREVUNAKKKKAATVV SSEEDHLKAEARGDMSSEALLILDEFTILARDL YAFYPLLIREVUNAKKKKKAATVV SSEEDHLKAEARGDMSSEALLILDEFTILARDL ALANRPSIKDKINTSNINGKREGONFVYQNEINN MSSILITDTKSKMSKAVAVQNEINN MSSILITDTKSKMSKAVAVQNEINNILGGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFFILEGVSKWIETEEHYPELOGLELD ARVWHKLLSKQRRAVVACFRMAPLYNLPR HAVNILFULGVSKSWIETEEHYPELOGLELD ARROWALLYKDLPNRTDDTSDPEKTVERVL DIANVLFFILEGKSKRVAVACFRMAPLYNLPR HAVNILFILGVSKSWIETEEHYPEROKLEDLA KROMSPERDEGGTKNVDPLHQLILLISRTALI							
YLAGGLQSCOMLVSKGYPDIGWNPVEGER YLDFLRAFVENGGSVANVVVNLIRRPE CFGPALRGEGONGLLAAMEEAIKIAEDPSRD GPSPNSGSSKTLDTEEEEDDTHIMGANIMTFY SALIDLGGAPEMHLHAGKGEARIRSILRS LIPLGDLVGVISIAFQMPTIAKDGMYVEPDMS AGFCPDIKAAMVLFLDRYYGIEVQDFLJHLL EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTRAQRDSIEVCLLSICGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH YERCWKYVCLPGGWGFGASEEELHLSIRK LFWGFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMGSSYVSNMCSSMDSEGORIPNQ PVDTSNITHEKLEYFINKYAEHSHDKWSMDK LANGWYGEIVSDSKVQPLMKPYKLLSEKE KEIYRWPKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMEAMMSYYNNIWAKKKKM ELESKGGGNIPHLVPYDTLTAKEKAKDREKA QDILKFLONGYAVSRGLELDTPSIEKRPA YSFLQQLRYVDEAHGYILFPDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGGIASNKEEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTYMKTG LESVKSALRAFLDNAAEDLEKTMENLKQOGF THITNOPKGVTQINYTTVALIPMLSSLFEHI GQHGFGEDLLEDVGYRLTSLYGGTSK SIYVERQRSALGECLAAFAGAPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGRYTQMPIIVMEVILPM LCSYMSRWWEHOPENNPERAEMCCTALNSE HMTTLLGONILKIIVNIGJDEGAWMKKLAVF SOPIINKVRPQLLKTHFLPLMEKLKKAATVV SEEDHLKAEARGOMSCKILTHPLPLMEKLKKAATVV SEEDHLKAEARGOMSCALLINEFITLADL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVALSPITYWSKSINFTRREEDNFVQNEINN MSPLITDTKSKMSKAAVSDQERKKMKKGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LAKNFFSLKUTELEGVSRKWKGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LAKNFFSLKUTELEGVSRKWETEHTYFEDKLEEDA KROUPHER DER SKRUNGLEDP AIRWQMALYKOLPNRTDDTSDPEKTVERVL DIANVLFHELGYSKE WETEHTYFEDKLEEDLA KAVNIFLGQSFERSWETEHTYFEDKLEEDLA							
YLDFLRFAVPCNOGSSVERANVVVRLLIRREE CFGPALRGEGONGLIAAMFEAKIAEDPSRD GPSPNSOSSKTLDTEEEDDTHMGNAIMTTY SALIDLLGRCAFEMHLHAGKGEARIRSILRS LIPLGDLVOVISIAFOMPTIAKDGNVVEPDMS AGFCPDHKAAMVLFLDRVYGIEGEARIRSILRS LIPLGDLVOVISIAFOMPTIAKDGNVVEPDMS AGFCPDHKAAMVLFLDRVYGIEVODFLLHLL EVGELPDLRAASLDTAALSATDMALALNEY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSIGGQLRPS MMQHLLRRLVFDVPLLNEHAKMPIKLLTNH YERCWKYYCLPGGWGNFGAASEELHLSRK LFWGIFDALSQKKYEGEHLALPCLSAVAG ALPPDYMESNEYVSMMEKGSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSSK VQFLMKPYKLLSEKE KEIYRWPIKESLKIMLARTMRTERTREGDSM ALVYNTRTRISOTSQVSVAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNIPILLVPYDTLTAKEKAKDRSKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSSLQQLIRYYDEAHQYLLEFDGSSRGKGEHI PYEQEKFFAKVVLPLIDQVYKNHRLYFLSAA SRPICSGGHASNKEKEMVTSLFCKLGVLVNH RISLFGNDATSIVNCLHIGQTLDATVMKTG LESVKSALRAFLDNAAEDLEKTMENKGGF THITNQPKGYTQINYTTVALLFMLSSLFEHI GQHQFGEDLLEDVQVSCYRLTSLYALGTSK SIYVERQRSALGECLAAFAGAPVAFLETHLD KHNIYSIVTKSSERFAALSLPTIVEDVCPNIP SLEKLMEEIVELAESGIRYTOMPITVMEVILPM LCSYMSRWEHGPENNPERAEMCCTALNSE HMNTLLGNILKITYNLGIDEGAWMKRLAVF SQPIINKVRPQLLKTHFLPLMEKLKKAAATVV SQEIDHLKARARGDMSALSLIDHTFTLADDL YAFYPLLIRFDYNRAKWLKEPNPEAEELFR MVALSFITDTKSKMSKAAVSDQERKKMKRKGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LANNFSLKDTEDEVSURGRHYCLVEHPPORSK KAVWHKLLSKQRRHYCLVEHPORSK KAVWHKLLSKQRRHYCLVEHPORSK KAVWHKLLSKQRKRAVACFRMAPLYNLPR HRANNIFLQGYEKSWIETEHTYFEDKLEDLA					4		I
CFGPALRGEGGNGILAAMEEAIKIAEDPSRD GPSPNSGSSKTLDTEEEDDTHIMGNAIMTPY SALIDLLGRCAPEMHLHIAGKGEARIRSILRS LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS AGFCPDHKAAMVLFLDRYYGIEVQDFLJHLL EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH YERCWKYVCLPGWGFGAASEELHLISK LFWGFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSSKVQPLMKFYKLLSEKE KEIYRWPIFESLKTMLARIMRTERTREGDSM ALYNRTRRISGTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGRIPPLLVFYDTLTAKEKAKDREKA QDILKFLQNOYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYSGGIKFFAKVVLPLIDQVFKNHRLYFLSAA SRPLCSGGASNKEKEMVTSLFCHGVVHRH RISLFGNDATSINVCHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THTRNOPKGVTQINITYALIPMLSSLFEHI GQHOFGEDLILEDVQVSCVRILTSLYALGTSK STYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSTYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEETVELAESGIRYTQMITIVETVILDM LCSYMSRWEHOPENDFRAEMCCTALNSE HMNTLLGNILKINYNJLGIDEGAWMKRLAVF SQPIINKVRPQLKTHIFLKKKKAKATVV SEEDHLKAEARGMSBAELLILDEFTILARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVABVFIYWSKSINTKERGOFNYVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMGTSLIVAALKRLLPIGLNICAFGQGELIA LAKNRFSLKDTEDVSTQRINHLQKLEPD AIRWQMALYKOLPNRTDTSDSPKTVERVI. DIANVLFHLEQKSKR VQRRHYCLVEHPORSK KAVWFIKLLSKORKRAVACFRMAPLYNLPR HRAVNLFLLQGYEKSWIETEEHYFEDKLIEDLA							` `
GPSPNSGSSKILDTEEEEDDTHIMGNAIMTFY SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS LIPLGDLVGVISIAFQMPTIAKDGNVYEPDMS AGFCPPHKAAMVLFLDRVYGIEVQDFLJHLL EVGFLPDLRAAGSLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLDSLLHT VYRLSKGCSLTKAQNSEVCLLSIGGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH YERCKGYVCLPGGWGNFGAASEELHLSRK LFWGIPDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFNKAEHSIDDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTEGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMAENYHNIWAKKKM ELESKGGGNIPILLVPYDTLTAKEKAKDREA QDILKFLQNGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYYDEAHQYLIDGYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSINVACHIGGTLDARTVMKTG LESVKSGGNIPILLVPYDSCYRILTSLYALGTSK SRYVERQRSALGECLAAFAGAFVAFLEHDL GQHQFGBDLLEDVGVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFVAFLEHDL KHNIYSIVTKISSRERAALSLPTNVEBUCPNIP SLEKLMEEIVELAESGIRYTQMIYIVMEVILPM LCSYMSRWEHOPENNPERABMCCTALISE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKAAATVV SCEEDHLKAEARGOMSEGNIFVGWITIVMEVILPM LCSYMSRWEHOPENNPERABMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SCEEDHLKAEARGOMSEGONFVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLINICAFGOQELIA LAKNRTSLKDTEDEVSREONFVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLINICAFGOQELIA LAKNRTSLKDTEDEVSREONFVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLINICAFGOQELIA LAKNRTSLKDTEDEVSREONFVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLINICAFGOQELIA LAKNRTSLKDTEDEVSREONFVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAFGOQELIA LAKNRTSLKDTEDEVSREONFVQNEINN MSFLITDTKSKMSKAAVSCRRMYCLVEHPORSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEHTYCLUEHPORSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEHTYCLUEHPORSK KAVWHKLLSKQRKREVGLUEHPLGLILLESTTALT							
LIPLGDLVGVISLAPQMPTIAKDGNVVEPDMS AGFCPDHKAAMVLFLDRVYGIEVQDFLHLL EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLITRAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKGCSLTKAPLFAGTEHHASLIDSILHT VYRLSKAPLFAGTEHHASLIDSILHT KENTYPERESILKTMLARTMRITERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSSVAPLMARTERTREGDSM ALYNRITRISGTSTAPLAST VSEQGIANNEKEMYTSLFCKLGVLVR RISLFGNDATSTVACHTAREA VSELOGLAPAGTSVARLTSLAFLAR VSELOGLAPAGTSVARLTSLYALGTSK STYLERGRAPLSTULLEFTSLAFLAT GQPHGFGBULLEDVQVSCYRILTSLYALGTSK STYLERGRAPLSTULLEFTSLAFLAT GQPHGFGBULLEDVQVSCYRILTSLYALGTSK STYLERGRAPLSTULLEFTSLAFLAT GQPHGFGBULLEDVQVSCYRILTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSK STYLERGRAPLSTULTSLYALGTSC STYLERGRAPLSTULTSLYALGTSC STYLERGRAPLSTULTSLYALGTSC STYLERGRAPLSTULTSLYALGTSC STYLERGRAPLSTULTSLYALGTSC STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYALGT STYLERGRAPLSTULTSLYA							l .
AGFCPDHKAAMVLFLDRVYGIEVQDFLJHIL EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMHLRRLVFDYPLNEHAKMPLKLLTNH YERCWKYYCLPGGWGNFGAASEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWYGEIYOSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMAENYHNIWAKKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDRRKA QDLKFLQJNGYAVSRGFKDLELDTPSIEKRPA YSFLQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRICZSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQFKGVTQINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRATISLYALGTSK SIVYERGRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLFTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVRVPQLLKTPLJMEKLKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFFYLLIRFVDVNRAK WLKEPNPEAEELFR MVAEVFIYWSSSIHPLJMEKLKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFFYLLIRFVDVNRAK WLKEPNPEAEELFR MVAEVFIYWSSSIHFLJMEKLKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFFYLLIRFVDVNRAK WLKEPNPEAEELFR MVAEVFIYWSSSIHREDONINCONDEINA MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLIPIGLNICAPGDQELIA LAKNRFSLKOTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLEFHLEGKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGVEKSWIETEEHYFEDKLIEDLA							SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRLVFDVPLLINEHAKMPLKLLTNH YERCWKYYCLPGGWGNFGAASEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNTIPEKLEYFINNYAEHSDDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISGTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKM ELESKGGGNHPLYVDTDTTAKEKAKDREKA QDILKFLQNGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHGYJHEPDGGSRGKGEHF PYEOEIKFFAKVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGOF THITRNOPKGVTQINYTTVALLPMLSSLFEHH GOHQFGEDLIL EDVGVSCYRLTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTINVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPINKVKPQLLKTHFLPLMEKLKKKAATVV SGEDHLKAARGMSEAELLILDEFTTLARDL YAFYPLLIRFVOYMAKWLEPNPEAEELFR MVAEVFIYWSKSINFKREGONFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRGGD RYSMGVTSLIVAAKLPLFIGLNICAPGDGELJA LAKNRFSLKDTEDEVROPINSNIHLQGKLEDP AIRWOMALYKDLPRINDTSDPEKTVERVL DIANVLFHLEGKSKWGRAHYCLVEHPORSK KAVWHKLLSKORKRAVACFRMAPLYNLPR HRAVNLFSLEGGTIKRVDPLHQLILLFSKTALT							,
LCTAVLPLATRCAPLFAGTEHHASLIDSLLHT VYRLSKGSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRRLVPDVPLLNEHAKMPLKLLTNH YERCWKYYCLPGGWGNFGAASEEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESN YSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINFYAEHSHDKWSMDK LANGWYGEHYSDSSK VQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYIHNIWAKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGMDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGOF THITRNQPKGVTQINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRLITSLYALGTSK SIYVERQRSALGECLAAFAGAPPVAFLETHLD KHNIYSYNTKSSERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENPPERAEMCCTALNSE HMNTLLGNILKINNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAABAGGDMSEALLLDEFTILADL YAFFYLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEGKSKRVGRRHYCLVEHPQRSK KAVWHKLLSK QRKRAVVACFRMAPLYNLPR HRAVNLFLSK QRKRAVVACFRMAPLYNLPR HRAVNLFLSK QRKRAVVACFRMAPLYNLPR HRAVNLFLSK QRKRAVVACFRMAPLYNLPR HRAVNLFLSK QRKRAVVACFRMAPLYNLPR HRAVNLFLSK QRKRAVVACFRMAPLYNLPR				:			
VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH YERCWKYYCLPGWGNFGAASEEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTIMFTERTEEGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKM ELESKGGGNIPLLVPYDTLTAKEKANDREKA QDILKFLQNGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMYTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGGYILDATIVNKTG LESVKSALRAFLDNAAEDLEKTMENLKQOF THITRNOPKGVTQIINYTIVALLPMLSSLFEHI GOHGFGEDLILEDVQVSCYRILTSLYALGTSK STYVERQRSALGECLAAFAGAEPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTINVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVÜLPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHILKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFYDVNRAK WILKEPNPEAEELFR MVAEVFIYWSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMGTSLIVAKKRLPLGILICAPGDGELIA LAKNFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANNLFHLEGSKRVGRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGKRKVDPLHQLILLFSRTALT					ŀ		
MMQHLLRRLVFDVPLLNEHAKMPLKILTNH YERCWKYYCLPGGWGNFGAASEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKM ELESKGGNIPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSGFKDLELDTPSIERRPA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAAEDLEKTMENLKQQGF THITRNOPKGYTQIINYTTVALILPMLSSLFEHH GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNYSIYNTKSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM LCSYMSRWEHGPENNPERAEMCCTALNSE HMWTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHIFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILLDEFTTLARDL YAFYPLLIRFYDYNRAK WLKEPNPEAEELFR MVAEVFIYWSKSHNFKREGONFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEGKSKRVGRRHYCLVEHPQRSK KAYWHKLLSKQRRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVOPLHQLILLFSRTALT	. I				j		
LFWGFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKY AEHSHDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALVNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQVILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMYTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGGTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQOGF THITRNOPKGVTQINYTTVALLPMLSSLFEHI GQHOFGEDLILEDVQVSCYRLTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKITVNNLGIDEGAWMKRLAVF SOPIINKVRPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRRVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINFKREQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMGTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQKYEKSWIETEEHYFEDKLIEDLA KFGAEPPEEDGTIKRVDPLHQLILLISRTALT			1	İ	İ		l
ALPPDYMESN YVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAEMYHNIWAKKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYYDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMYTSLFCKLGVLVRH RISLFGNDATSINVCLHLGQFLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THTRNQPKGVTQINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMFILYMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINFKREQNFVVQNEINN MSFLITDTISKSMSKAAVSDQERKKMKRGGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEBEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVCRVL DIANVLFHLEGKSKRVGRRHYCLVEHPQRSK KAVWHKILSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQKYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTIKRVDPLHQLILLEFSRTALT			1		ĺ	ſ	
PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNIPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGSSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THTRNOPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAPPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMTIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHEPLPMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDVROTIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA							
LANGWIYGEIYSDSSK VQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTIMRTERTREGDSM ALYNRTRISQTSQVSVDANHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKM ELESKGGGNIPILVPYDTI.TAKEKAKDREKA QDILKFLQNGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQVFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCVRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLJTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMFIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFYDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVDDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA							-
KEIYRWPIKESLKTMLARTMRTERTREGDSM ALYNRTRRISQTSQVSVDAHGYSPRAIDMS NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQELKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKENPPEAEELFR MVAEVFIYWSKSINFKREEONFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWIKLLSKQRKAAVVACFRMAPLYNLPR HRAVNIFLQGYEKSWIETEEHYFEDKLIEDLA				ĺ			
NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNIPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSSRGGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGYTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMFIIYMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	1 1		- 1				
ELESKGGONHPLLVPYDTLTAKEKAKDREKA QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQLIRYVDEAHQYILEFDGGSRGKEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGOILDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGGF THITRNQPKGVTQINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHILD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAFLLILDEFTTLARDL YAFYPLLIRFVDYNRAK WLKEPNPEAEELFR MVAEVFIYWSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKR VGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT							ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF THITRNQPKGVTQINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHILD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTILLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINIFKREGONFVVQNEINN MSFLITTDTKSKMSKAAVSDQERKKMKRGD RYSMQTSLIVAALKRLLPIGINICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT							= -
YSFLQQLİRYVDEAHQYILEFDGGSRGKGEHİF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMYTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGVTQIINYTTVALLPMLSSLFEHİ GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPHVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMMTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			1				
PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIY WSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	}	}	1	1			
SRPLCSGGHASNKEKEMYTSLFCKLGVLVRH RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKGGOF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALINSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP AIRWQMALYKDLPNTTDDTSDPEKTVERVL DIANVLFHLEQKSKKVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNIFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	1	Ī					
LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIY WSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			Í				
THTRNQPKGVTQIINYTTVALLPMLSSLFEHI GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPHIVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT							RISLFGNDATSIVNCLHILGQTLDARTVMKTG
GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPHVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVV SQPIINKVKPQLLKTHFLPLMEKLKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT							
SIŸVÈRQRSALGECLAAFAGAFPVAFLETHLD KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIR YTQMPHVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAK WLKEPNPEAEELFR MVAEVFIY WSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT				ļ			
KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP SLEKLMEEIVELAESGIRYTQMPHVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			- 1	}			
SLEKLMEEIVELAESGIRYTQMPHVMEVILPM LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			İ	l			
LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			i	ļ			SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM
SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSINFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT							LCSYMSRWWEHGPENNPERAEMCCTALNSE
SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	1	1	ľ	i	1	ŀ	
YAFYPLLIRFVDYNRAK WLKEPNPEAEELFR MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT					İ		` `
MVAEVFIYWSKSHNFKREEQNFVVQNEINN MSFLJTDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLJVAALKRLLPIGLNICAPGDQELJA LAKNRFSLKDTEDEVRDJIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYPEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT		Į		ļ	į	ľ	
MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			ļ	ŀ	į		
RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	•	ļ	j	j	į	†	
AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT					ŀ		RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			ŀ			1	•
KAVWHKLLSKQRKRAVVACFRMAPLYNLPR HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT			l			1	
HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT		ŀ	l			[K V WHKI I SK OBK B A V V A CEDMADI VAII DD
KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT	[[- 1	ĺ	- 1	İ	ļ	HRAVNLFLOGYEKSWIETEEHYFEDKLIEDI A
		ŀ		1	ļ	j	

NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide seq- uence NO: of nucl- octide location op/496 O	Histidine, ne, Proline, ine, tophan,
eotide sequence USSN location corresponding to last amino acid residue of peptide sequence USSN 100	ne, -Proline, ine, tophan, o codon,
sequence 09/496 corresponding to last amino acid residue of peptide sequence 09/496 corresponding to last amino acid residue of peptide sequence 09/496 corresponding to last amino acid residue of peptide sequence M=Methionine, N=Asparagine, P=Q=Glutamine, R=Arginine, S=Ser T=Threonine, V=Valine, W=Tryptosine, X=Unknown, *=Stop /=possible nucleotide deletion, \=p nucleotide insertion EEEVKSFEEKEMEKQKLLYQQ	-Proline, ine, tophan, o codon,
uence 914 ng to first amino acid residue of peptide residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 926 937 737 74 75 75 75 75 75 75 75 75 7	ine, tophan, codon,
amino acid of peptide T=Threonine, V=Valine, W=Tryp residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop peptide sequence nucleotide deletion, \=p nucleotide insertion EEEVKSFEEKEMEKQKLLYQQ	tophan, codon,
residue of peptide /=possible nucleotide deletion, \=p nucleotide insertion residue of peptide /=possible nucleotide deletion, \=p nucleotide insertion EEEVKSFEEKEMEKQKLLYQQ	codon,
peptide /=possible nucleotide deletion, \=p nucleotide insertion EEEVKSFEEKEMEKQKLLYQQ	
sequence nucleotide insertion BEEVKSFEEKEMEKQKLLYQQ	OSSIDIE
EEEVKSFEEKEMEKQKLLYQQ	
	ABI UDBCAA
NSTVOOKMLDYLKEKKDVGF	
CSVLDLNAFERQNKAEGLGM'	
LODDEFTCDLFRFLOLLCEGH	
QTGNNTTVNIIISTVDYLLRVQ	
SGKDVIDEQGQRNFSKAIQVA	1
QGPCTGNQQSLAHSRLWDAV	
QMKLSQDSSQIELLKELMDLQ	
MLEGNVVNGTIGKOMVDMLV	· ·
KFFDMFLKLKDLTSSDTFKEY	
RDFHKAMESHKHYTQSETEFL	
TLDYEEFVKRFHEPAKDIGFN	
MPNDTRLQTFLELAESVLNYF	
SAKRIERVYFEISESSRTQWEK	· .
FDVVNEGGEKEKMELFVNFCE	
QISESDLNERSANKEESEKERP	
SILTVRSALFALRYNILTLMRM	
KKVKKMTVKDMVTAFFSSYW	
ASVFRGFFRIICSLLLGGSLVEG	i i
ANMPDPTQDEVRGDGEEGERI	KPLEAALPSED
LTDLKELTEESDLLSDIFGLDLI	KREGGQYKLIP
HNPNAGLSDLMSNPVPMPEVO	EKFQEQKAK
EEEKEETKSEPEKAEGEDO	GEKEEKAKED
KGKQKLRQLHTHRYGEPEVPE KGKQKLRQLHTHRYGEPEVPE	SAFWKKIIAY
QQKI,LNYFARNFYNMRMLAL	FVAFAINFILL
FYKVSTSSVVEGKELPTRSSSE	NAKVTSLDSS
SHRIIAVHYVLEESSGYMEPTV	1
TEQPSEDDIKGQWDRLVINTQS	
VKRKVMDKYGEFYGRDRISEL	
FSDAREKKKPKKDSSLSAVLN	SIDVKYQMW
KLGVVFTDNSFLYLAWYMT	67777770000000000000000000000000000000
553 1903 A 4199 31 767 LPELNGRGAGLRRAEPSERGG	• • • •
LPLSHGHSHGGGGCRCAAER/ AYGLYLRIDKGRLQCLNESRE	
1	
ERAD\DRSKFVESDADEELLFN GIIIMGEDDDSHPSEMRLYKND	,
EPDQ'TFSLNRDLTGELEYATKI	٠ .
HISKNFGADTTKVFYIGLRGEV	
CNYEASANPADHRVHOVTPOT	
554 1904 A 4200 1 961 GIPCTEMGNFDNANVTGEIEFA	
EICIKACKNLAYGEEKKKKCN	
RSSQGKRKTGVQRNTVDPTFQ	
OLVTROLOVSVWHLGTLARRY	
WDFEDSTTQSFRWHPLRAKAL	
NGELTVRAKLVLPSRTRKLQE	
HGQLCLVVLGAKNLPVRPDGT	` `
LPDQQKLRLKSPVLRKQACPQ	
TPAQLRQSSLELTVWDQALFG	
RLGSKGDTAVGGDACSQSKLQ	
LWTDMTLVLH	
555 1905 A 4211 331 2419 KENKKARNLRMNQSRSRSDGG	
NHHENERRWQQERLHREEAY	
DYRLMRDHNLLGTPGEITSEEL	
QLASQPDLRDGTNYRDSEVPR	ESSHEDSLLE
WLNTFRRTGNATRSGQNGNQT	
NNGEFRFSLEIHVNHENRGFEI	
DSNRDHTANRQQRST\SPVARF	
GSSSNIPRTRLASRGQNPAEGS	FSTLGRLRNGI

uence 93/496 correspondi 1914 and to first and to get decidence ca	SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
peptide	seq-			09/496	correspondi ng to first amino acid	to last amino acid residue of peptide	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
GQRFGAAHVWENGARSNYTYRNTNQRLEPI RINTSINSRSSRSPIGGTYVHNOGRSPRY QQTTRRSVERRGRTRVFLEQDRERERGTAY TPSNSRLVSRITVEGEGESSRSTA VSRIPTIT LDLQVRRIPPGENDGNDSIANRTRSRVGLAE ATVTESNSGFPETISRLERSGIFTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MGADSESELQRNQCHLPDMSESINLGTDN NRSQHREGSSQDRQAQDSTENHEGENETTYQ HTRNSDSRGGRQLRNPNNLVETGTLFILRLAH FELLNSSDDDDRRGLTREQIBLYSTHYEHEN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HHCIDRWLSENCTCHCRQPVLOSNIANNG GARGE LORQRQIPAAPAYVRCFTECTDDVIMPRR SSPENTERGKGGSVVTKQFFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGGSVVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGGSVVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGGSVVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGSVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGSSVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGSSVTKGFTRSARLASAPA PPKPEPKPRXTSAKKEPGAKISRGAKCKKEEK KSPENTERGKGTASENGETKAEHHERSTVNSTS SRGTPPSTLSVKGQIETVRVKGTEN ARTSCLTLGTSWGHRFGPPRANNVVLVET GPLHIGQAGHKLPTSGDPPAASQSARITGMS HTWPLASFLIDSCKTHYVEINYTIMS STATISTAKASAPANVYSCTTETT GPLHIGQAGHKLPTSGDPPAASQSARITGMS HTWPLASFLIDSCKTHYVKIMYTIMS TYRHAEREPHETSSATKVSTYYRRIKEPKLLD GOOPDSDGRTQKYCKEBKKYSEGKGLAKE LOCPTITICRGRETQDGQVKEPFKPSKKDIJAC TYSKKDYDULRSSNBVWEKKKKEGHOCKE SNSSSNQLDKSSKLPDVKPSPRINLRKKSLTTVK VDVKKTTUDFRVF1GEFAQEIT HIMQVKANYPPSGGTTHERRSKAADHHADV NEPLNSDPHHRRIBNSLALGSKQAVIVESS QTILKITIDPHOLRIFIDERRKERLONGEDHIFH ASAAERDQNSSSTNYTTORKOITHKPFEV EGNIRNTRVEPKSNTTIGKGRCQPNYKSGLVQ KSLTQAKYQKLEATRAGAGTGVS HHAWLQVLNS SSEGGGLGGISHCLISERSERVEPVN GGGTSEDLFWKLDALGTPRICHFERLIPP LOLDSPHALGGSRHAGAGTGVS HHAWLQVLNS TSSEGGRHRADARAWSDLMVEH ARTSCLTTURGGGRHRALSCRPVINGEFREP LIMDPLRTGLLCGNCKFHKHLDDLPAPLVV RAMUDLOGSSPHAGHRGLISTERSERVEPVN GGGTSEDLFWKLDALGTRAFTHAETEVKLDT KREMTLVARFVTLECHAANSONDEDEDEED SELVQVILDKOKKEPWIRKLEDTUR SFLSTTVKAASKVVDVPPPGGMVADAVVTF VRHSQDVLDKTNBERMTERLFQCWYNSSM NVICTWLTDRAMDLQLHTVGLAKLSRYDGGTFS SFLSTTVKAASK					peptide	Sequence	/=possible nucleotide deletion, \=possible
OQTTRRSYRRGETRYFLEQDRERERGTS VARRIPTT							GQRFGAAHVWENGARSNVTVRNTNQRLEPI
LDLQVRRIRPGENRORDSIANETRSRVELS							QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
MEADSSELQKNGGHLPDMHSELSNLGTDN NRSODRAGQGDSDRAGQGDSTGMGHENTETOP HTRNSDSRGRQLRNPNNLVETGTLPILR.IAH FFILIMESDDDDRIRGLTKEQIDNLSTRHYEIN SIDSELGKICSVCISDYVTGNKIRQLPCMHEF HIHICIDRWLSENCTCPICRQPVLGSNIANNG GRORGHPAAPAVPVRCFTCFTDIVIMPKR KSPENTEGKDGSKVTKQEPTRESARLSAKPA PREPEPRENTSAKEPGAKISGAKGKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKQGIPTVRVKGTSN SKSTEPSTLSVKGGIPTVRVKGTSN SKSTEPSTLSVKGGIPTVRVKGTSN SKSTEPSTLSVKGGIPTVRVKGTSN SKSTEPSTLSVKGGIPTVRVKGTSN SKSTEPSTLSVKGGIPTVRVSQKGPLNRE LDCFNTGRGRETQDQVKEFFKPSKKDGIPT LDCFNTGRGRETQDQVKEFFKPSKKDGEKE SNSSSNQLDKSQKLPDVKPSPINLRKSLTVK VDVVKTVDTFRVASSYTERGMSLDLVAVG RKSENFHVFEHLDSTQNTENKFTGFFAQEIIT IIHQVKANYPFSSGTTLBERFSKMADIHKADV NEIPLNSDEFIHRRIDMSLAELGSKQAVIVESS QTILKIIDPNDLRIDDERRKSELTVK VDVVKTVDTFRVASSYTERGKGCQPNVKSGLVQ KSLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP KSLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP SKLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP SKLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP SKLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP SKLVQAKVQRLRTTGPRGFITHKFRERLMK KKVP SKLVQAKVQRLRTGPRFAVWSDLMVEN LALVIVSKLSVAKIGKKREMYELPVCLQ ALVNYSRLSVAKIGKKREMYELPVCLQ ALVNYSRLSVAKIGKKREMYELPVCLQ KLCSMEMGGGFAKMWHQHDAVTT KEMITLJVAKVTVILGULARLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHVELEGTGK HLEQRILKMASDMBESCVKRTKRIJAFEVKLQK HLEQRILKMASDMBESCVKRTKRIJAFEVKLQK SLSQGGGGAKMWHQHDADAVTT VRISQDVLRDKVNEEMYERLFDQWYNSSM NVICTWLTDRMDLQLHTVQLKTLIRMVKTY RSPRILQVLDSTLINSKTYSTIRNRLTVEEDATA SVSEGGGLQGISMKDSDEEDEEDD SSEGGGLQGISMKDSDEEDEEDD SSEGGGLQGISMKDSDEEDEEDD SSEGGGLQGISMKDSDEEDEEDD SSEGGGLQGISMCKSDEEDEED							LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE
HTRNSDSRGGRQLRNPNNLVETGTPLILLA.H FFILLNESDDDRIBGLITEQUIDM.STRHYEIN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLOSNIANNO 356							LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN
SIDSELGKICSVCISDYVTGNKLRQLPCMHEF							HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH
1906							SIDSELGKICSVCISDYVTGNKLRQLPCMHEF
O. O. O. O. O. O. O. O.	556	1906	A	4212	3	462	LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR
1907 A 4213 774 507 ARRESCLTLQTSWGHRENGPPREVARIVELVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFIVYKIMYTL 558							QEAGKEGTAPSENGETKAEEIHISRSTVNVST
HRTWFLASFLIDSCKNFIVYKIMY*TL	557	1907	Α	4213	774	507	ARRESCLTLQTSWGHRH\GPPRP\ANFVFLVET
GDQPFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWEEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFIHPVFFHLDSTQNTENKPTGEFAQEIIT IIHQVKANYPFSPGITLHERFSKMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNIRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 559 1909 A 4235 1 323 KFSIPPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGPLIET EFHHVQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLN 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNOKDAENVGRLTPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPDTVDSFPLFQ LLNDFLRTGLLICGNGKUFHKHLQDLFAPLVV RYMWDLDGSSPIAQSHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEFGK HLEQRLKLMADDMESCVKRTRIKAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLVSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHTYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKPIKRADILKHVIGDYKDI	550	1000		4225		1252	HRTWFLASFLIDSCKNFIVYKIMYTL
TYSNKNDVDLRSSNDKWKEKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKSLITVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGTLHERFSKMADIHKAADV NEIPLNSDPEHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDITHKPPEV EGNHRNTRVRPFKSNFRGGCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 559 1909 A 4235 I 323 KFSIPFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVQQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TISLEERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLINDFLRTGLLICGNGKFHKHLQDLFAPLVV R/YMWDLDGSSPIAQSHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVRRTRIIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KLICSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRNSQDVLRDKVNEEMYTERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	338	1908	A	4223	3	1255	GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE
RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYPFSPGITLHERFSKMADHIKADV NEIPLNSDPEHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAERDQNSSFSKNYTTQRDDITHKPPEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 559 1909 A 4235 1 323 KFSIPFFLRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGPLIET EFHHVQQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVMDAALEVQPPDTWDSFPLFQ LLNDFIRTGLLICGNGKPHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRIAFEVKLQK TSSIQQIFRVPGFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKLDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE
NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTILKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPPEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 1 323 KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 1 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVPCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWSDLMVEH AETFISLFAVDMDAALEVQPPDTWDSFPLFQ LLINDFLRTGLLICGNGKVHKHLQDLFAPLVV RYMWDLDGSSPIAQSHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRIKLMASDMESCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKFGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	-						RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT
ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 559 1909 A 4235 I 323 KFSIPFFLRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLNDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTRIIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL/CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERJFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE
KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP 559 1909 A 4235 1 323 KFSIPFFLRWSFTLVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGPLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ ILINDFLRTGLLICGNGKVFHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPKRADILKHVIGDYKDI							ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV
LGLSHSPASASQVGGITGTQHHTGLIFGPLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK
HHAWLQVLNS 560 1910 A 4246 2 1569 TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGK/FHKHLQDLFAPLVV RYYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR/IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	559	1909	A	4235	1	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL
AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGKVFHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR/LAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							HHAWLQVLNS
LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI	560	1910	A	4246	2	1569	AALVNYSRLSEYAKIEGKKREMYELPVFCLA
LLINDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							LAELVIEVLQQNEEHHAEAFAWWSDLMVEH
GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV
TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK
KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI					•		TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP
VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS
RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD 561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							VRHSQDVLRDKVNEEMYIERLFDQWYNSSM
561 1911 A 4257 1300 654 SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI							RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA
	561	1911	A	4257	1300	654	

NO of mucl-todide sold	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence USSN docation corresponding of periodic sequence unice of the periodic sequence of periodic se			•				
Sequence	1		1 1100	1		1	D-Aspartic Acid, g-Giutanic Acid,
Sequence	1	,					r=Pnenylalanine, G=Glycine, H=Histidine,
1914 ng to first amino acid residue of peptide peptide peptide sequence 1-Threnine, V-valia, W-tryptophan, Y-vyosine, X-Uaknown, "-Siop codon, P-possible nucleotide deletion, P-possible nucleotid	cotide	seq-		l .	_		
ammo acid residue of peptide residue of peptide residue of peptide sequence Fi-Threonine, V-Valine, W-Tryotine, W-Sop codon, /-possible nucleotide deletion, V-possible nucleotide deletio	seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ammo acid residue of peptide residue of peptide residue of peptide sequence Fi-Threonine, V-Valine, W-Tryotine, W-Sop codon, /-possible nucleotide deletion, V-possible nucleotide deletio	иепсе	ł	İ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine.
Peptide Pept		į					
Popsible nucleotide deletion, \(\) possible nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide nucleo		ļ		ĺ	l .		
	l	l .		i		sequence	
Intlepuedaemgodgiptiglimiusgi	l .		ļ				
FMKGNTIKETEAWDELLALIGVYPTKKHLIFG DPKKLITEDFYRGPYLIPDVJEFQ WGRTNLETSKMKVLKFVAKVHNODPKDW PAGVCEALADEENRARPOSOPADS S62	ŀ				sequence		nucleotide insertion
FMKGNTIKETEAWDELLALIGVYPTKKHLIFG DPKKLITEDFYRGPYLIPDVJEFQ WGRTNLETSKMKVLKFVAKVHNODPKDW PAGVCEALADEENRARPOSOPADS S62							INTLEPVEEDAEMRGDOGTPTTGLLMIVLGLI
						ĺ	
	1	ŀ		ļ		1	
PAQYCEALADEERRARPQPSGPAPSS	1		ļ	ì	J		, ,
1912 A 4260 1 1498 MVTWLYRFLPTSNMAAKIRSLLPPDIRUDG WHARLOKCELSROCGSYCAGASPLOK WARMCHMICGREELLRILLOGGRRYHSVAGPSO WHARLOKCELSROCGSYCAGASPLOK WARMCHMICGREELLRILOGGRRYHSVAGPSO WHARLOKCELSROCGSYCAGASPLOK WARMCHMICGREELLRILOGGRRYHSVAGPSO WHARLOKCELSROCGSYCAGACHON WHARLOKCELSROW WARROW WHARLOKCELSROW WARROW WHARLOKCELSROW WHARLOKCELSROW WARROW WHARLOKCELSROW WARROW WARROW WHARLOKCELSROW WAR	1	ł					
WILHARLOKCELSRGCGSYCAGAKASPLPUS WAGNALMGGRRELRILDGSGRYNSVAGPSQ WLGKPLTTRLIFPAAPCCCRPHYLSVAGPSQ WLGKPLTTRLIFPAAPCCCRPHYLSLAASGPR SLSTSAISFAEVQVQAPPVVAATSPJTAVPEV ASGETADVQTAAEQSFAELGLGSYTTYGLI QNLLEFMIVDLGPWGAGAACTVFAKCI PLIVITGGREAARINHLPEIGKFSSKREAKLA GDHIEY YKASSEMALYQKKHGIKLYKPLIVE TQAPPISSFIALREMANLPYPSLGTGGLWP QDLTVSDPJYILPLAVTATIWAVLELGACT VQSSDLQWMRNVIRMPLITLIPTIMIFPTAV FMYWLSSNLFSLVQVSCLRPAWRTVLKIPQR VVHDLDKLPPREGFLESKKGWARAMTQ VLEREQRAWRIQLELARGPLRQTFTINPLQ PGKDMPPMPSSSSSSSKFKSKYMAEMTQ LERERQRAWRQLELARGPLRQTFTINPLQ PGKDMPPMPSSSSSSSSKFKSKYMHOTIG WOMEN PROSINPPMPSSSSSSSKFKSKYMHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VVHDLDKLPPREGFLESKTKGWHOTIG VHDLDKLPPREGFLESKTKGHTTSML VHDL			1	l		ļ	PAQYCEALADEENRARPQPSGPAPSS
WIHARLOKCH-ISROCGSYCAGAKASPLPUS WARMALMGGRREILRILQSGRRYNSVAGPSQ WLCKPLTTRLIFPAAPCCCRPHYLFLAASGPSQ WLCKPLTTRLIFPAAPCCCRPHYLFLAASGPSQ WLCKPLTTRLIFPAAPCCCRPHYLFLAASGPSQ SLSTSAISFAEVQUAPPVVAATSPTAYPEV SAGETADVVQTAAEQSFAELGLGSYTTYGLI QMLLEFMIVDLGLPWGGAIAACTYFARCH PLIVTGQREAARENHHLPEIQKFSSKREAKLA GDHIEY YKASSEMAL YQKKHGIKL YKPLIUCH QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSLQWMRNIVIRMPLILTIPTHIFFTAY FMYWLSSNLFSLVQVSCLRPAWRTYLKIPQR VVHDLDKLPPREGFLESKKGWARAMTRQ LERRQRARNIQLELAARGPLRQITTTINPLIQ PKDMPPMIPSSSSSSKEKSKYMAEMTRQ LERRQRARNIQLELAARGPLRQITTTINPLIQ PKDMPPMIPSSSSSSKEKSKYMAEMTRQ LERRQRARNIQLELAARGPLRQITTTINPLIQ PKDMPPMIPSSSSSSKEKSKYMAEMTRQ LERRQRARNIQLELAARGPLRQITTTINPLIQ PKDMPPMIPSSSSSSKEKSKYMHDTIG WGGLAPTGTLEPTIREYQNTQLSYYLLPEPMSFQPVGGVAPAQWATHTINPLIQ PROSINPHOTELIATAPPALLPHPTHTPQSF LIQENNTINHTHSHTHITYTETLSFFLYICVNN DRIMEWGKSVF LIQENNTINHTHSHTHITYTETLSFFLYICVNN DRIMEWGKSVF LIQENNTINHTHSHTHITYTETLSFFLYICVNN DRIMEWGKSVF LIQENNTINHTHSHTHITYTETLSFFLYICVNN DRINEWGKSVF PROSINCHARLOR PRO	562	1912	A	4260	1	1498	MVTWLYRFLPTSNMAAKLRSLLPPDLRLOF
MAMGLMCGRRELLRLLQSGRRYHSVAGPSQ WLGKPLTTRLLPFAAPCCCRPHYLFLAASOFR SLSTSAISFAEVQVQAPPVVAATRSPTAVPEV ASGETADVVQTAAEQSFAELGLGSYTPVGLI QNLIEPMHYDLGLFWWGALAACTYFARCLIF PLIVTGGREAARIHNILPEIQKFSSRIREAKLA GDHHEYYKASSEMAL YQKKHGIKLYKPLLPV TQAPIFISFFIALREMANLPVPSLQTGGLWWF QDLTVSDPYILPLAVTATMWAVLELGAETG VQSSDLQWMRNVIRMMPLITLPTIMHFFTAV FWYMLSSNLFSLVQVSCLRFAVTVLKIPQR VVHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VVHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VVHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VVHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VVHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNQLELAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNGLEAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNAUELAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNAUELAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ LRREGRMRNAUELAARGPATVLKIPQR VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRQ VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKKGWNAEMTRG VHIDLDKLPPREGFLESFKGWNAEMTRG VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDLDKLPPREGFRAIGHT VHIDL							,
WIGKPLITRILEPAAPCCCRPHYLELAASGPR SISTASISAEVQVQAPPVAATSPTAVPEV ASGETADVVQTAAPCOYOAPPVAATSPTAVPEV ASGETADVVQTAAPCOYOAPPVAATSPTAVPEV ASGETADVVQTAAPCOSPAELGLGSYTPVCLI QNILEPMPVDLGLEWWGAIAACTVFARCLIF PLIVTGQREAARHINIEPIGKPSSRREAKLA GDHIEY YKASSEMAL YQKKHGIKL YKPLILPV TQAPHISIFFILAREMANIPVETGGELWWP QDLTVSDPIYILPLAVTATMWAVLELGAETG VQSSDLQWMRVIRMMPLITLITIMHPFTAV PMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR VVHDLDKLIPPREGPLESFKKGWKNAEMTRQ LRREGRMRNQLELAARCPLRQTFTTINPLLQ PGKDNPPNIPSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSSSSKYKSKYPWHDITLD PGKDNPPNIPSSSSSTRYBSQPVGCK HRSPNGGLFRGSPVK/TPIPMSFQPVPGGVL PRGSGPPHGTSILTAPPALLPHPPTHPTQSF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTELSFFLYICVNN DRMEWGRSVF LIQENNITHHFISHTHYTENSFFLICHTMLP PAWILIMICQAPROPOPPRITOPERSSLOEAF GOTGASTPET CHARLES AND AND AND AND AND AND AND AND AND AND	1			1	Ì		
SISTSAISPAEVQVQAPPVVAATPSITAVPEV	ł	ł		i	ł	ł	
ASGETADVYQTAAEQSPAELGLGSYTPYGLI	i						
		1	l		1	1	ASGETADVVQTAAEQSFAELGLGSYTPVGLI
FLIVTOGREAABINHIPEIQKESSRIREAKLA GDHEYYKASSEMAL VQKSSILQWF QDLTVSDPIYLIPLAVTATMWAVLELGAETG VQSDLQWMRNVIRMIPITINTHPTTAV FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR VVHDLDKLPPREGFLESFKQWKNAEMTRQ LREREQRMRNOLBLAARGPLRQTFTHNPLLQ PGKDNPPNIPSSISSSSSKPKKYPWHDTLG FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR VVHDLDKLPPREGFLESFKGWKNAEMTRQ LREREQRMRNOLBLAARGPLRQTFTHNPLLQ PGKDNPPNIPSSISSSSKPKKYPWHDTLG FMSTNGGLAPTQTLEPTNEYQPYGGVL PROSGNPPHGTSILTAPPALLPHPTHPTQQSF LIQENNTHNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF FIGURIANTHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF S64 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLQVGPAAQWAYLFTINSLQGF FIFLVYCLLSQQVQKQYQKWFREIVKSKSES ETYTLSSKMOPDSKPSEGDVPPRTSE S65 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELGTMLP PAWLLMLQAPROPDPPRLTOPEKSLQEAP GQTGASRITRT S66 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNNSTVA DHFDHRAVIGSEFICIGNDTGTGRFPGGLE DVSTYPVLIEBLLSRSWSEEELQGVLRGHLLR VFRQVEK VREESRAQSPKEPFYGGLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLOSVGGTGTSDTGWGALRCQ MIFPAGALVCSRLGRWBTQKRCPDPSYFS VLNAFIDRKDSYYSHQLAQMGVCGCKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHAMD NTVYMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRRLVLLIPLLGLTD NEAPVETLKHCFMMRQSLGVIGGKPNSAHY FIGYVGEELIVLDPHTTQRAVEFGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SREFLITTIPLYLTJTSFYTKYDQHFVLTVVS	1	l	l		l		
GDHIEYYKASSEMALYQKÄGIKLYKPLILPY TQAPIFISFIALREMANLPYPSLQTGGLWYF QDLTVSDPIYLPLAVTATMWAVLELGAETG VQSSDLQWMRNVIRMMPLITLPITIMFTTAV FMYWLSSNLSILQVSCKLIPAVRTVIKIPQR VYHDLDKLPPREGFLESFKGWKMAEMTRQ LREREQRMRNQLELAARGPLRQTFTINIPLLQ PGKDNPPNIPSSISSSSKKSKYPWHDTLG S63	j .		1]	ļ	J	
TQAPITISFFIALRAMAID-PPSLQGGLWWF	1 '		I	l			
	1	Ì	[l			
VQSSDLQWMRNVIRMMPLITLPITMHPFTAV	ļ	1	Ī	1	1		
VQSSDLQWMRNVIRMMPLITLPITMHPFTAV	Į l	ļ		I	1]	QDLTVSDPIYILPLAVTATMWAVLELGAETG
FMYWLSSNLFSLVQVSCLRIPAYRTVLKIPQR]			
VYHOLDKL PREGFLESFKKGWKNAEMTRQ LREREQRMRNQLELAARGPLRQTFTHNPLLQ PGKDNPPNIPSSSSSKFKSKYPWHDILG PGKDNPPNIPSSSSSKFKSKYPWHDILG PGKDNPPNIPSSSSSKFKSKYPWHDILG PGKDNPPNIPSSSSSKFKSKYPWHDILG PGKDNPPNIPSSSSSKFKKYPWHDILG PGKDNPPNIPLIPSSATMPSMQCK HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGVUL PRGSGNPPHGTSILTAPPALLPHPPTHPTQGSF LIQENNINHHISHTHITYTETLSFPLYICVNN DRMEWGKSVF DGKTWCLGILQVGPAQVMAYLFTIINSLQGF FIFLVYCLLSIQQVQKQYQKWFREIVKSKSES ETYTLSSKMGPDSKYPEGDVFPRTSE PFYALLATTKPALGTIRLYPCSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQFEKSLQEAP GQTGASRTPRT LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLITLTFTCSTPWAESSTKFHHMYTNVS GLTSFGEKVVEELNRLGMMDLSYASDTLIRR VLEVSQAPVIFSHSARAVCDNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPGQLE DVSTYVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEPPYGQLSTSCH FHLGASEWTPRLIJWR FIGYVGELIJWR SATPLYPTORT SATP		Î	1	1	i		
LREREGURMNOLELAARCPLRQITTHIPPLLQ	1		į.	i '			
PGKDNPPNIPSSSSSSSKKKSKYPWHDTLG							,
1913 A 4265 623 116 MGGLAPTQTLEPTREYQNTQLSVSYLIPEQN THGTRRTLSSGPSNNLPPLSSSATMPSMQCK HRSPNGLFRGSPVK/TPPIMSFQPVGGVL PRGSGNPPHGTSILTAPPALLPHPPTHPTQGSF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTTLSFFLYICVNN DRMEWGKSVF LIQENNITNHTHSHTHTYTTLSFFLYCOPPERING LICENSE LIQENNITNHTHSHTHTYTTLSFFLYCOPPERING LICENSE LIQENNITH LIQENSE LIQENNITH LIQENSE LIQENNITH LIQENSE LIQENNITH LIQENSE LIQEN	1			1			
THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK HRSPNGGLFRQSPVKTPPFMSFQPVFGGVL PRGSGNPPHGTSILTAPPALLPHPPTHPTQGSF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF ILKKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLLQVQPAAQWMAYLFTINSLQGF FIFLVYCLLSQQVQKQVQKWFREIVKSKSES ETYTTLSSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLAYECSPELGTIMLP PAWLIMLCQAPRQDPDPRLTQFEKSLQEAP GQTGASRTPRT LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHIMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDCTGREPPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNPAIGGTGPTSDTGWGCMLRCQQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVPDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGITD INEAYVETLKHCFMMPQSLGVIGGRPNSAHY FIGYVGEELIYLDPHTTQPAVEPTIGCFIPDES FHICQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITIPVLYFILTSFYTKYDQHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	i i		ł	i	i		PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK HRSPNGGLFRQSPVKTPPFMSFQPVFGGVL PRGSGNPPHGTSILTAPPALLPHPPTHPTQGSF LIQENNITNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF ILKKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLLQVQPAAQWMAYLFTINSLQGF FIFLVYCLLSQQVQKQVQKWFREIVKSKSES ETYTTLSSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLAYECSPELGTIMLP PAWLIMLCQAPRQDPDPRLTQFEKSLQEAP GQTGASRTPRT LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHIMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDCTGREPPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNPAIGGTGPTSDTGWGCMLRCQQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVPDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGITD INEAYVETLKHCFMMPQSLGVIGGRPNSAHY FIGYVGEELIYLDPHTTQPAVEPTIGCFIPDES FHICQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITIPVLYFILTSFYTKYDQHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	563	1913	A	4265	623	116	MGGLAPTOTLEPT\REYONTOLSVSYLLPEON
HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGVLL	505	17.5	٠.	1203	023	***	
PRGSGNPPHGTSILTAPPALLPHPTHPTQQSF	1						
LIQENNNTNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF 564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLŁQVGPAAQVMAYLFTIINSLQGF FIFLVYCLLSIQQVQKQYQKWFEIVKSKSES ETYTLSSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQFRQAFVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELGTTMLP PAWLLML.CQAPRPQDPDPRLTQPEKSLQEAP GQTGASRIPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVYLTLTIFTCSTPWAESSTKFRHIMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGFTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQLAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRGLTD INEAYVETLKHCRMMPQSLGVIGGKPNSAHY FIGYVGEGLIVLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLITTIPLYFLTSFYTKYDQHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY				<u>'</u>			
DRMEWGKSVF			l	,			
DRMEWGKSVF				i I			LIQENNNTNHTHSHTHTYTETLSFFLYICVNN
1914	1	1	1				
GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF FIFLVYCLLSIQQVQKQYQKWFREIVKSKSES ETYTLSKMGPDSKPSEGDVFPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELGTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRIPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNILNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHIPDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVPDTWSSLAVHLAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPVLLIPLRIGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTOPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQHFVLNTVS	564	1014	Ā	4270	2	360	
FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES ETYTLSSKMGPDSKPSEGDVPPRTSE 565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTILRLYECSPELGTTMLP PAWLLML.CQAPRPQDPDRRLTQPEKSLQEAP GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRPPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVPDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFEREFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQHFFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1 304	1714	Ι ^ .	4270	,	300	
565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFPHIRAVIGSEFIGIGGNYDGTGRFPQGLIE DVSTYPVLIELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASSEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGTGKMDAATLTYDTLRFAEFE DPFTSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGFTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHJAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTTFIVLYFLTSSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1		l				
1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELGTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT	i i		1	1			FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
PPSALLAPTKPRALGTLRLYECSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITIPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	i						ETYTLSSKMGPDSKPSEGDVFPRTSE
PPSALLAPTKPRALGTLRLYECSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITIPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	565	1915	A	4288	83	406	RNSRPLWCSPPASOPROAPVSOSCCCPLPSSSS
PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMYTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTTTPIVLYFLTSFYTKYDQHFFVLNTVS LMSVLIPKLPQLHGVRIFGINKY		17.15	**	-250	-		
GQTGASRTPRT 566 1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQHFFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1		[
1916 A 4298 1041 229 LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL'E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIVLDPHTTQPAVEFIDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568				-			
GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY							
GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLANVSDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWLIGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	566	1916	Α	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL'E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY			1				GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLVE DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRIGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1 1		ł	i i			
LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGFTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRIGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY			l				· ·
DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\(\text{V}\) DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY			l				
DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1		Ì				
DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY							DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E
VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	j !						
FHLGASEWTPRLLIWR 567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	j l						
567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVT:NRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY]						
DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1-60	10:-	<u> </u>	4000	L.——	1106	
WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRLFGINKY	j 567	1917	Α	4299	1	1106	
WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRLFGINKY							DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRILGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRLFGINKY	j l						
VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRILGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRLFGINKY	j l						
WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRLFGINKY	į l					[
NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1						
CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	1						WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY							NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY) i	•					T I
FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY]						
FHCQHPPCRMSIAELDPŠIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	[
GAECCLGMTRKTFGFLRFFFSMLG 568 1918 A 4300 2012 1843 SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	į į			١ ٠			
568 1918 A 4300 2012 1843 SRKFLTTTPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	į l						FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
568 1918 A 4300 2012 1843 SRKFLTTTPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY	; l						GAECCLGMTRKTFGFLRFFFSMLG
LMSVLIPKLPQLHGVRIFGINKY	568	1918	A	4300	2012	1843	
			**				
1919 A 4302 186 531 WTFCLFL/WWVPESARWLLTQGHVKEAHRY	<u>- </u>	1010		4202			
		צועו	А	4302	180	251	WITCLTL/WWVPESAKWLLIQGHVKEAHKY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI
						CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT VTESKLEAEGKTKEKAREKERKKKS
570	1920	A	4308	3		RSGQGKVYGLIGRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	Α	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYV\ETL\KHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLK\ERM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV TQGTLFIILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	Α	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	**	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
		'		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]		l		peptide	Jouque	/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
577	1927	Α	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
						ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
					į	FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	Α	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
ŀ		İ				SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
]						SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
ļ.		ľ	!		}	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
						CWPGWSSTPDLK
580	1930	Α	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
		l				\VFKKGI\IHILHELFQNKEEGAFPNS/FYEASFT
			•			LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
		·	j			QLKSSDL
581	1931	Α	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
1						RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
		ŀ				RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
í í		ĺ				RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
i						LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
		l				VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
						DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
			ļ			SPE
582	1932	Α	4424	194	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
1 1				,	•	LEQELLEHGRDAASVQAATSVQAMQGKTTL
1	1000	L				PS\QGPLQRPSRLVFT\DVANAIHV
583	1933 .	Α	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
						PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
584	1934	A	4439	1	628	SAPPALLQDTSV
364	1734	A	4439	'	028	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ APPQQLPSAQPPQPSNPPHGAHTLNSGPOPGT
l i		İ				APATQHSQAGPATGQAYGPHTYTEPAKPKK
						GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
						ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
						GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
]				ļ		QGPHGKAAQGGAAGAAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
					• • •	SIFDDFAHYEKRQ
586	1936	Α	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
						FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
						INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
						TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
						FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
					ŀ	LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
			}	1		PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
						ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
		.				LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
	İ			ľ		PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
	Ī			-		LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
			[l		SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
500	1027		440:			FRAPPAINARLPFNFFFPFV
587	1937	Α	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
	j	J	j	}	J	CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
500	1029		4400	1720	1450	NKHEPPCPATKRRHSGSI
588	1938	Α	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
	I	l		l	Į	CPANFCIII/DFLVETGFHHVGQASHELLTSGD
589	1939	,	4407	022	222	PPTSASQSAGITGMSYHTWFGES
207	1737	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
1		Į	į			PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
	ļ		j		1	REAAQAMYIDSYNSRGFMINGNRVLGPCALL PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI
	i			!		11104 1 A MILL GOLIÁDITE DOLOFIL METEL KIEL

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN 09/496	location	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	uerice		914	correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence			1717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
Ì		ĺ	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
Į				peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
						DTPNACATFNFLCHEGRVTGAALIPPPGGTSL
590	1940		4492		472	TSLGQAAQ
390	1940	A	4492	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
		ł	ł		ł	RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
		į	l		ì	ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
	ŀ					VPHQGGLPGPIRVAPSSAGQREASQGPPGR
591	1941	Α	4495	1444	1116	IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT
						MEYYADTERNEIMSF\AGTWVELEAIILSKLM
						LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL
	1045	<u> </u>				EPWDSSCFPHPSSGV
592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
		l				VCGGLSLLANAWGILSVGAKQKKWKPLEFL LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
		Ì				EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
						MWMVCWPVNYRLSNAKKQAGHTVMGIWM
						GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
					ŀ	GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
]		ļ	j			AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI
Ī						DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL
593	1943		4506		102	GPFSLADTHLSDLPYTWGDRDSGGACVM
393	1943	Α	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC HFPASASQVAGTTHARHHTQLIF\AFLVENGL
	i					C
594	1944	A	4507	1327	647	KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
						ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
]						CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII
	ĺ		[-		NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ
						RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
						GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG
						PADGSLYI
595	1945	A	4512	533	264	FFFKMESYSVARLECSGAISAPCNLHLLGSNN
Ī						SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
						YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	Α	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD
					-	HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
						NLMLELLVQKKKQLEAESHAAQLQILMEFLK
						VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPOFEAPSPSHSSIIDSTEYSOP
,						PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
						DLEOCYFSTRMSRISDDSRTASOLDEFOECULS
						KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV
			l			FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
						VDIHYPENEMTCNSKISCISWSSYHKNLLASS
						DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
						DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
						ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
						EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
						EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
						KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
						CWRALPDGESNVLIAANS\QGTI\KVLELV
597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS
						ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE
600	1040		450		204	QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
598	1948	Α	4524	1	384	FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF
			لـــــا			TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID NO: of NO: of NO: of nucl- eotide seq- uence uence	-
nucl- eotide seq- uence	
eotide seq- uence USSN location corresponding I=lsoleucine, K=Lysine, L=Leucine, seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, P=Pro	
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, P=Pro	dine,
	line
	mic,
amino acid of peptide T=Threonine, W=Tryptoph	120
residue of sequence Y=Tyrosine, X=Unknown, *=Stop coo	
peptide /=possible nucleotide deletion, \=possi	
sequence nucleotide insertion	010
RRREMQSQSVMLALRRGDAVWL	SHDHDG
YGAYSNHGKYITFSGFLVYPDLAF	AAPPGLG
ASELL	7201020
599 1949 A 4526 366 776 MGOPAPYAEGPIOGGDAGELCKC	DFLVFTSP
NPEAVCEAGTPAMFOTAWROME	
VQWRDPGSLHPPPLGFKRFSCLSL	•
HAPPHPANFCIFSRDQVSPCWPGW	
PPPWLPKVLGLQA	011022211
600 1950 A 4529 776 334 FFFETESCYVAQAGVQWCDLCSL	OAPPPGISS
DPPASASRVAGTTGARHHTQLIFV	
\MLARDGLKLLTSSDPPASASOSS\	
RLANFFVFLVETGSRYVAQAGVO	
LLISTGYLTCSYSDLGRFTPP	
601 1951 A 4533 1460 403 HEVOESIHFLESEFSRGISDNYTLA	LITYALSS
VGSPKAKEALNMLTWRAEQEGGI	
SKLSDSWQPRSLDIEVAAYALLSH	
GIPIMRWLSRQRNSLGGFASTQDT	
EFAALMNTERTNIQVTVTGPSSPS	
HNRLLLQTAELADGTANGSV/SIS/	
CQLNVYNVKASGSSRRRRSIQNG	
AVKENKDDLNHVDLNVCTSFSGP	•
MEVNLLSGFMVPSEAISLSETVKK	
LNLYLDSVNETQFCVNIPAVRNFK	
SVSIVDYYEPRRQAVRSYNSEVKL	
VQRLPSL	
602 1952 A 4540 1963 295 MRAPGRPALRPLPLLLLLLSSP	WGRAVPC
VSGGLPKPANITFLSINMKNVLQW	TPPEGLQG
VKVTYTVQYFIYGQKKWLNKSEC	RNINRTYC
DLSAETSDYEHQYYAKVKAIWGT	KCSKWAE
SGRFYPFLETQIGPPEVALTTDEKS	ISVVLTAP
EKWKRNPEDLPVSMQQIYSNLKY	NVSVLNT
KSNRTWSQCVTNHTLVLTW\LEPN	1TLYCVHV
ESFVPGPPRRAQPSEKQCARTLKD	QSSEFKAK
	'IHVG\KEK
HP\ANLILIYG\NEFDKRFFVPA\EK	V\INFI\TL
NIS\DDSKISHQDMSLLGKSSDVSS	
NLRPPQEEEEVKHLGYASHLMEIF	
CAGPEEQELSLQEEVSTQGTLLESC	
GPQTLQYSYTPQLQDLDPLAQEHT	
EPSTTLVDWDPQTGRLCIPSLSSFD	
PSEGDGLGEEGLLSRLYEEPAPDR	PPGENETY
LMQFMEEWGLYVQMEN	
603 1953 A 4543 3 600 YSAVEFVEQASGISDWWNPALRK	
GMIAPYYEDSDLKDLSHSRVLQSP	
LQAVIAGDLMKLIESYKNGGSLLIG	
LHYAAETGNGEIVKYILDHGPSEL	
TGETALHKAACQRNRAVCQLLVD	
TDSKGKTPQERAQQA\GDPDLAA/	YTIESRQN
YKVIGHEDLETAV	
604 1954 A 4548 3 938 QDNKVQNGSLHQKDTVHDNDFE	
QSNSYPSMSDPYLSSYYPPSIGFPY	
STAGDPPIPYLTTYGQLSNGDHHFI	
QPGGLGNNIYQHRFNFFPENPAFSA	
QGQQTQSSAYGSSYTYPPSSLGGT	
FHSDTLSKAPGMNSLEQGMVGLK	1
VKTVGSVVSSVALTGVLSGNGGTI	
KPTSWAAIASKPAKPQPKMKTKSO LPPPPIKHNMDIGTWDNKGPVPKA	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	<u> </u>			sequence		nucleotide insertion
						SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ QPPQ
605	1955	Α	4553	2	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC
ļ	1	1				ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS
	ŀ	•	,			FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS
	ł					SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
						GPGLGALDRLRAHASAMGDEDLPGMAALQP
	[ł			1	HGVPGDGEGPHERGPPPASAPVGGTVTLRED
		l				SAKRLERRARRISACLSDYSLASDSGVFEPLT
	ł	į]	KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG
1						SECLLVHVLOLKNPAGLAVKEDCKVHIRVYL
	1	l				PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
1	}	1			}	HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
1]				LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
		l				LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
i i	ł	ł	}			TTAQLQAVERELAEERAKLEYTEEEVLEMER
1		l				KEEQAEA!SERSWQADSVDSGCSNCTQTSPPY
	1					PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
1		ŀ				KVDKETNTEDLFLEEAASLVKERPSRRARGSP
1		ļ				FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS
		1				STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL
						MARTSLDLELDLQASRTRQRQLNEELCALRE
1 1	ĺ	ĺ				LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
		1				EAERQTRQTKLDYRHEQAAEKMLKKASKEI
		}				YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL PADDV
606	1956	Α	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
		ļ				VLLLQDSSGDYSLAHVREMACSIVDQKFPEC
(VEDEQUOSODISE ATTREMACTIVE QUAITEC
1						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRILSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEML WGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGINLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWELAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSSSFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV
		!				GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEML WGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL
		!				GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVITYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563	1	4499	GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKHHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	Α	4563	1	4499	GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRISNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	Α	4563	1	4499	GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSSSFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI
607	1957	A	4563	1	4499	GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRRISNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETETEMKALGERVSIL SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
scq-	uence	J	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	İ	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		}	}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	· ·	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		peptide		/=possible nucleotide deletion, \=possible
<u> </u>		L ·] .	sequence	ļ	nucleotide insertion
						HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
		į .			Ì	VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
			{			DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
1			ĺ			TPEELNCERPRITSEPQDADVTSGNTVYFTCR
		į				AEGNPKPEIJWLRNNNELSMKTDSRLNLLDD
	ĺ					GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
1 1						EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
j .						ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS
						GGLYIQNVVQGDSGEYACSATNNIDSVHATA
						FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
1						GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
i l						SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ
						PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
		-				EPAITWNKDGVQVTESGKFHISPEGFLTINDV
						GPADAGRYECVARNTIGSASVSMVLSVNVPD
1						VSRNGDPFVATSIVEAIATVDRAINSTRTHLF
						DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
1						ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
,						PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY
						RTHDGTCNNLQHPMWGASLTAFERLLKSVY
						ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
						GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
						VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
						DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
i .						VYPREQINQLTSYIDASNVYGSTEHEARSIRD
1 1						LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
ŀ						RDENESPIPCFLAGDHRANEQLGLTSMHTLW
						FREHNRIATELLKLNPHWDGDTIYYETRKIVG
1 1			i i			AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
j j				-		PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD
						ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
						RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
1						VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
						AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
i i						LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR
						DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
						NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
1						LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
						FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS
			ľ			A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR
]						TQIKKLESR\LSTTECVDAGGESHANNTKWK
				[KDACTICECKDGQVTCFVEACPPATCAVPVNI
					_	PGACCPVCLQKRAEEKP
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
ļ İ				. 1		IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
[· l		GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF
						YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ
		. I		l	l	LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD
				l		EFTIAATDDNHIFAWGNGGNGRLAMTPTERP
) <u> </u>						HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
1 1				ļ	i	VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
ļ <u>.</u>]				GGPDAW
609	1959	A	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
į į			1		ŀ	PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
					ļ	HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
	İ		1	İ		RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
L						KLSFHVLSGS
610	1960	Α	4570	697	467	ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
		1	l			GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N
<u> </u>						LVIRPPRPPKVLGLQA

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	ĺ	09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
nence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	'	/=possible nucleotide deletion, \=possible
<u> </u>				sequence		nucleotide insertion
611	1961	Α	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
į						WNPNVPESPRIPAPRLPKRMSGAPTAGAALM LCAATAVLLSAOGGPVOSKSPRFASWDEMN
	İ	ĺ	1	İ	ĺ	VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
						LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
	1	Ì				LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
	{	ĺ				RIQHLQSQFGLLDHKHLDHEVAKPARRKRLP
						EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
-						QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
					ļ	EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
		ĺ	İ			VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
				: 		LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
						GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
612	1962	A	4575	162	3	WRGRYYPLQATTMLIQPMAAEAAS FFFETESRSVAOAGVOWRDLSSLOPPPPG\SR
012	1902	A	43/3	102)	GSPASASPVAGITGTRHHRTRG
613	1963	A	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
						SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
ļ	1					HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
						LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	Α	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
		ſ				GGLFCAWVGTILLVVAMATDHWMQYRLSGS FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
						GIFYMCAYRVHECRRLSTPR
615	1965	Α	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
ĺ		[ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
						PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT
		}				TRLPAGSQEPVKD
616	1966	A	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
	İ	<u> </u>				SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
	1055		128.			AGLELPTSGDPPALASQSAGITGVTTVPSGPG
617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
}						DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKOLOELOOOOEEEEROROORREERROONL
						RARSREHPVVGHPDPALPPSGVNCSGCGAEL
}						HCQDAR*
618	1968	A	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
ļ						SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
ļ					,	MVGNETTYEDGHGSRKNITDLVEGAKKANG
ļ						VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
İ					•	HCYMEYSRLRGEAGSDVSLVDLGFOTDFRV
						YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI
]						AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
						AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
					[CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG
,						AFASYYWALRKPDDLPAFPLFSAFGRALRYH
						TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
1						KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
ł						IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
1 .						TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
						APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
						LNKTNKKAAES
619	1969	Α	4601	2	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY
		ــــــــــــــــــــــــــــــــــــــ		<u> </u>		2

NO of Month and Selection (1997) and Selection (199	SEQ ID	SEO ID	Mat	LEEO	Deadistad	I Designation of	
nucleotide sequence uence use sequence per corresponding to corresponding to the sequence per corresponding			Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence Sequence		1	noa	L .			D=Aspartic Acid, E=Glutamic Acid,
1970 1970 A 19		1 * *		1			
1914 1915 1916							
minio acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence pertia sequence peptide s		uence	l	l .			
Peptide Peptide Sequence Peptide Sequence Peptide Pe	uence	ŀ		914			
Peptide Sequence	j	ļ		}			
	l .			l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
GLAYGHESYESNIRDW-DULGGW-VTGNOK	i	İ		l	peptide		/=possible nucleotide deletion, \=possible
GLAYGHESYESNIKDVVDLQGWYTGNOK GIVILIDPOHISVOQKVFTINGKGKGIFYEN GOLVILIDPOHISVOQKVFTINGKGKGYESTEN GOLVILIDPOHISVOQKVFTINGKGKGYESTEN MCHURCHEICHBLSLTRPSMEKPCKS MCHURCHEICHBLSLTRPSMEKPCKS MCHURCH FORGOLD SPRSSGTYVQKVEGPR KGHI LEEEEDGEGAETLAHFCPMELRGPEP LCSRRPOPLIPWAAAGRAAPILY LYTALIJF TGAFLLGYVAFRGSCQACGSVIAVVSEDVAN YEPDLDFHQGRLYWSDLQAMFLQFLGEGRI EDTIRQTSLRRRVAGSAGMAALTQDIRALIS ROKLDHWITDTHYVGLQPPDPAHPNTLHWV DEAGK VOGEQHLEDDYVCPYSAIGHANTLELV DEAGK VOGEQHLEDDYVCPYSAIGHANTLELV VYAHYGRPEDLQDLRARGVDPVGRLLLLVRV OFFICE OF A STATE OF	1	1		í	sequence	Ĭ	nucleotide insertion
GLIVLTIDPQHISVDQKVFTINFGKRGIFYFFN NQHVECNEICHALSLITRSMEKPCKS NQHVECNEICHALSLITRSMEKPCKS NQHVECNEICHALSLITRSMEKPCKS NQHVECNEICHALSLITRSMEKPCKS A 4606 1 2415 MERLWGLPGRAQQLSPRSSGTYVQKVEGPR KGHLBEEEDGEGEGACGLOSVLVYSEDVR KGHLBEEEDGEGEGACGLOSVLVYSEDVR VERNER COSPROPNLIPWAAAGRRAAPYLVTALLIF TGAFLLGVVAFRGSCQACGDSVLVYSEDVR YPPDLDFHQGRLYWSDLQAMFLOPTGGEGRL EDTIROTSLRERVAGSAGMAALTQDIRAALS ROKLDH-WYDTHYVGLQPPDAHIPTLHW DEAGKVGEQLFLEPPDVYCFYSAIGRVTGEL VYAHYGREELQQDLRARGVDPVGRLLLVRV GVISFAQRVTHAQDFGAQGVLIYFEPADFSQ DPRFSLSSQQAVYGHHGTGDPYTGGFSF NQTQFPPVASSGLFSIRAQPISADIASRLLKIL KGPVAPQEWGGSLLGSYPHLGPGBPTLLVVN NIRTSTPINNIFGCTEGRSEPDHYVVIGAQRDA WGFGAAKSAYGTALLELVRTFSSWYSNGFR PRRSLLFISWDGGDFGSSYSTEWLEGYLSVL HLKAVVYSLDNAVLGDDKFHAKTSPLLTSL BESVLKQVDSRNISGGTLYEQVVFTNPSWDJ AFVIRLPMDSSANYSFTAVGVPAVEFSME DDQAYPFLHTKEDTYENLHKVLQGRLPAVA AFVIRLPMDSSANYSFTAVGVPAVEFSME DDQAYPFLHTKEDTYENLHKVLQGRLPAVA AFVIRLPMDSSANYSFTAVGVPAVEFSME DDQAYPFLHTKEDTYENLHKVLQGRLPAVA AVAOLAGQLLIRLSINGELLPLDFGFGYGDVV LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGDLKARGLTLQWYSARGDY LRHIGNLNEFSGLKARGLTLQWYSARGDY LRHIGNLNEFSGLKARGLTLQWYSARGDY LRHIGNLNEFSGLKARGLTLQWYSARGDY LRHIGNLNEFSGLKARGLTLQWYSARGDY LYMMINITAVAFFPALATILY LAFNLLGDGLADALDKKK SISVADPVGSGAANVINAVINAVAFFPALATILY LAFNLLGDGLADALDKKK SISVADPVGSGAANVINAVINAVAFFPALATILY LAFNLLGDGLADALDKKK SISVADPVGSGAANVINAVINAVAFFPALATILTY LAFNLLGDGLADALDKKKC LAYCRERVAVINAVINAVAFFPALATILTY LAFNLLGDGLADALDKKKTLINFK SCVILLGLLLLYVCRSFVYSVSVTVAVAVAYAY RNEDRWAWLQDLIGAGCLAULKTLINFK SCVILLGLLLLYVCRSFVYSVSVTVAVAVAYAY RNEDRWAWLQDLIGAGLAULKTLINFK SCVILLGLLLLYVCRSFVYSVSVTVAVAVAYAY LAFNLTGGARGANAVARIAVARGPALTVAQQFV AFVIRLYSVYSVTVAVAVAYAYAGGML TVVVFURMKFREVYS GGC							
1970 A 4606 1 2415 MERLWGLFGRAQQLSPRSSOTYLYVQRVEGER KGILLEEEEDGEGASTLAHCCPMELRGPEP LGSRFRQPNLIPWAAGRRAAPLVTQRVEGER KGILLEEEEDGEGASTLAHCCPMELRGPEP LGSRFRQPNLIPWAAGRRAAPLVTALLIF TGAFLLGY VAPRGSCQACCDSVLVVSEDVA YEPDLDFHQGRLWSDLQAMFLQFLGEGRL EDTIRQTSLRERVAGSAGMAALTQDIRAALTQDIPATLUVRV GVISFAQKVTINADDFGAQQVLIVPEPADFSQ DPPKPSLSSQQAVGHVHLGTODPYTGFSS DPFKPSLSSQQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTODPYTGFSS DPFKPSLSSQAVGHVHLGTGGAT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVGHVHLGT DPFKPLSSAVSAVSAVATALTED DPFKPLSSAVSAVSAVATALTED DPFKPLSSAVSAVSAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTED DPFKPLSSAVSAVSAVAVATALTA DPFKPLSSAVS	ł	ł	ł	1	l	}	GLIYI TOPOHSVOOK VETTNEGKRGIEVEEN
1970 A 4606 1 2415 MERLWGLFGRAQQLSPRSSQTYVQRVEGUE	ļ	1					NOHVECNEICHDI SI TERSMENECUS
KGHLBEEBEÖGEÄATILAHECYMELRÖPE LGSRPRQPNLIPWAAAGRRAAPYLVITALIJE LGSRPRQPNLIPWAAAGRRAAPYLVITALIJE LGSRPRQPNLIPWAAAGRRAAPYLVITALIJE LGSRPRQPNLIPWAAAGRRAATQDRAALS ROKLDHVWTDTHYVGLOPEPPAHPHILHW DEAGKVEGQILEDDYCPYSAIGNYTGEL VYAHYGREDLQDLRARGVDPYGRLLLYW GVISSAQKVTRAQDFGAQQULYPEPADFSQ DPPKSLSSQQAVYGHVHLGTGDPYTPGFPSY NQTQFPVASSGLSPIFAQFSSADIASRLLRKI KGPVAPQEWQGSLLGSPYHLGFGERILLVVN NIRISTEPNNIFGCIGERSEPDHYVGAQRDA WGFGAAKSAVGTAILLELVRITSSMYSIGFR PRSSLISIWDGGDFGSVGSTEWBGYSSAU HLKAVVYVSLDNAVLGDDKFHAKTSPLITSI LESVLKQVDSPNHSGQTLYRQVVYTANFSWD AEVIRPLPMDSSAYISTAFVGVPAVEFSFME DDQAYFFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAQQLILISHDRILLPGRYGDVV LRHIGNINEFSGDLKARGLTLQWVYSARGDV RAAEKLRGEIVSSEFDERLTRWTNVRIMRV EFYSLSQYVSPADSFFHIMGRGDHTIGAGL PHARLASVSSTOTGATSSTGFQGFSRFRRQL ALLITWDACKGAANALSGDVWHIDNIP SEYPLSQYVSPADSFFHIMGRGDHTIGAGL FYFLSQYVSPADSFFHIMGRGDHTIGAGL GSSIVYFFTMGTSTISASISTGLGGAQFTTP EWGAMLNEARADMVIAPHVAVPPALAFLTV LAFNLGDGLRADHVIKA GSSIVYFFTMGTSTISASISTGLGGAQFTTP EWGAMLNEARADMVIAPHVAVPPALAFLTV LAFNLGDGLRADHVIKA CGSSIVYFTMGTSTISASISTGLGGAQFTTP EWGAMLNEARADMVIAPHVAVPPALAFLTV LAFNLGDGLRADHVIKA SCYILLGDLICLLYDVFFVFITFFTIKNGESBWYEL AAGFFGNNEKNDONLVEATGGPSAPHEKLPV VIRVPRLIYFSVMSVCLMFVSSTYVSVTVAYAGML TVVLGUMKGQPALLYLVPCTLITACQFV VIRVPRLIYFSVMSVCLMFVSSTYVSVTVAYAGML TVVLGUMKGQPALLYLVPCTLITACQFV VIRVPRLIYFSVMSVCLMFVSSYPSVDSGQAG AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LLAYCREFDVGTGSSYTYVSVTVAYAGML TVVLGUMKGQPALLYLVPCTLITACQFV VIRVPRLIYFSVMSVCLMFLSSCLCAVERSAMMVIRSKY QYFDSRGMFISIVFSAPLLVNAMIVVNMVY KTLNVMTDLKNAGCRRKEKKERREDF GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI GGCCSGLNTMTMSAFSVSSPPVSIPATITQT DWTEFWLMKGPT TPCGHTYCTLCTINFLVERDCFCMDRKPLVL QHCKKSSILVNRLLNKLLIVTCFFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGGRARGE CLSECHHPTTSQAWGTHL*SQLLGGRARGE CLSECHHPTTSQAWGTHL*SQLLGGRARGE CLSECHHPTTSQAWGTHL*SQLLGGRARGE CLSECHHPTSCAWGAWGTHL*SQLLGGRARGE CLSECHPTHTSCAWGAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSECHTHOTTSCAWGTHL*SQLLGGRARGE CLSE	620	1970	ΙΔ-	4606	1	2/15	
LGSRPRQPNIJPWAAGRRAPYLVLTALLIF TGAFLIGVVAPRGCQACGDSVSSQVSSDVY YEPDLDFHQGRLWSDLQAMFLQFLGGEL EDTIRGTSLREVAGSAGMAALTQDIRAALS RQKLDHVWTDTHYVGLQFPDAFNTLHWV DEAGKVGGQLFLEDPDVYCPYSAIGNVTGL VYAHYGRFEDLQDLRARGVDPYGRILLLVRV GYISFAQKVTNQDFGAQVLIVPEPADDSS DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF NQTQFPPVASSGLFSIPAQPISADIASRLLIKL, KGPVAPQEWQGSLLGSPYHLGFGPRLLLVNN NIRTSTPRNIFGCEGRSSPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTSSWNSOFF PRESLLFISWDGGDFGSVGSTEWLGGVYSLV HIKAVYVSLDNAVLGDDEFHAKTSPLLTSL LGSVLKQVDSPNHSQQTLYEQVVFTNPSWD AEVIRPLINDSSAYSFTAFVGVPAVESFMD DDQAYFFHTHEGTTYENHKYLGGRLPAVA QAVQLAGQLIRLSHDRILPLDFGRYGDVY LRHIGINLRIFSGLKARGLILLQWYSARGDY IRAAGKLRGEIYSSERDERLTRMYNVRIMKY FEYELSQYVSPADSFPRHIMGRGDHTLGALL DHLRLLRSNSGTPGATSSTGFQGSSFRRQH FEYELSQYVSPADSFPRHIMGRGDHTLGALL DHLRLLRSNSGTPGATSSTGFQGSSFRRQH FEYELSQYVSPADSFPRHIMGRGDHTLGALL DHLRLLRSNSGTPGATSSTGFQGSSFRRQH GSSIVVFFTMRIGTSISAASLSFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV GSSIVVFFTMRIGTSISAASLSFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLDDGLRAALDHERPYGQ GTIACRGKMEVERUFILSGLCAVAVVWAVP RABENRAWNLQDILGJAFCLNLIKTLKLPRFYG GGSCCSGLNTMYPSAFSVSSPYVSIPAVLTQFT DWTERWLMGLATTHALCYLTCLSSRSYNGL AGPFGNNENDONLVEATOQPSAPHEKLPV VIRVPKLIYFSVWSCLMPVSILGGDIIVPGL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMIL LIAYCRFFDVQTGSSYTYVSVTVAYAIGMINVMWYLFSKY QFDSRGMFISIVFSAPLLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVAMMIVVNMWNLFSKY QFDSRGMFISIVFSAPLVAMMIVVNMWNLFSKY	020	13.0	1	7000	1	2413	
			1				
	ł	ľ	l	l		i	
EDTIROTISLEERVAGSAGMAALTÓPIRALS RÖKLDHWTDTHYGLOFPDPAHPNTLHWY DEAGKVGEQLPLEDPDYYCPYSAIGNYTGEL VYAHYGRPEDLOPLARGVDPVGRLILVRV GVISFAGKYTMAQDFGAGQVIYPEPADFSQ DPPKPSISSQQAYYGHYHLGTGDPYTPGFPSF NQTGFPPVASSGDISFIAQRISADIASRILRKI. KGPVAPQEWQGSLLGSPYHLGTGPPTTPGFPSF NQTGFPPVASSGDISFIAQRISADIASRILRKI. KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN NHRTSTPINNIFGCIEGRSEPDWEGYLSVI. HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSI. IESVIKQVDSPNHSGQTLYEQVVFTNIPSWD: AEVIRPLPMDSSANSFTAFVGPVAVEFSPME; DDQWYPFLHTKEDTYENHKVLQGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA QAVAQLAGQLIRLSHDRILKVLGRIPAVA GEYFLSQVSPADSPFRIIFMGRGDHTLGALL DHILLIRSNSSGTPGATSTGGVSFSFRROLD LILLILLSTSSEGEDREATERMYNVTIIMVS EFYFLSQVSPADSPFRIIFMGRGDHTLGALL DHILLIRSNSSGTPGATSTGGVSFSFRROLD ALLITWDACKGAANALSGDVWNIDNNF EFYFLSQVSPADSPFRIIFMGRGDHTLGALL HALLITWDACKGAANALSGDVWNIDNNF EFYFLSQVSPADSPFRIIFMGRGDHTLGALL LALTRUBACKGAANALSTUNAVFPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLLGGGLRADADVIRIVAVPPALAFILTV LAFINLGGGLRADADVIRIVAVPPALAFILTY LAFINLGGGLRADADVIRIVAVPPALAFILTV LAFINLGGGLRADADVIRIVAVPPALAFILTV VIRVPKLINTSVMSVCLMPVSILGFGDIDVPGL LIAVCRRPDVGTGSSTYTYVSVIVTAVAGMIL TVVVLGLIMKKGQPALLTVLYCTLITACQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRARDALTREDLAPSSVDSGQAGF GGCCSGLPNTMPSAFSVSSPPSUPAVLTOT DWFEPWLMGLATHALCVILLSSRSVRLQ LIAVCRRPDVGTGSSTYTYSVIVTAVAGMILTV WHWWWWW KTINNYTDLKNAQERNEKKKRRENED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LNNRGFI 624 1974 A 4622 164 668 VSCYTALOSIMNOPESANDPEPLCAVCGQAH SLEENHYSYPEVODDLICHICLQALLDPLD TPGGHTYCTLCLTITALVECDFCPMDRKPLVL QHCKSSLLVANKLLINKLLVTCFFREHCTOGQAH SLEENHYSYPEVODDLICHICLQALLDPLD TPGGHTYCTLCLTITALVECDFCPMDRKPLVL QHCK				Ì	i	İ	
RQKLDHVWTDTHYVGLQPFDPAHPNTLHWV	l	Į	į)	ļ		YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
RQKLDHVWTDTHYVGLQPFDPAHPNTLHWV	ľ		ł				EDTIRQTSLRERVAGSAGMAALTQDIRAALS
	ļ			1			
	1	l	1			ĺ	· -
GVISFAQIKVTNAQDFGAQGULIYPPADFSQ DPPKPSLSSQQAVYGHVHGTODPYTPGFPSF NQTQFPPVASSGLPSIPAQPISADIASRLIRKI KGPVAPQEWQGSLLGSPYHLGPQBIRLIVVN NHRTSTPINNIFGCIGGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSWISNGFF PRRSLIFISWDGDDFGSVEWLEGYLSVL HLKAVVYSLDNAVLGDDKFBLKTVSPLLTSI IESVLKQVDSPNHSGQTLYEQVLSVL AEVIRPIMDSSAYSITAFVGVPAVEFSPME DDQVAYPFLHTKEDTYEHLHKLQDGRLPAVA QAVAQLAGQLIRLSHBRITERMYNVRIMSD IESVLKQVDSPHSGQTLYEQVJSARGDV LRHIGKINFFSGDLKARGITLQWYYSARGDV RAAFELRGGYSSERREGETRMYNVRIMSP EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRILRSNSSGTPGATSSTGFQESSRRRQLA LATUNDACKGAANALSGDVMIDNNF EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRILRSNSSGTPGATSSTGFQESSRRQLA LATUNDACKGAANALSGDVMIDNNF STRVDDFVGSGJANVIIAVAITSSIPFARLVRGG NITULKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASI.SFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLDDGLRDADPKKG 622 1972 A 4614 2 820 LVYYMIAIFCIASAMSI.YNCLAALIHKIPGG CTIACRGKMWEVRLIFI.SGLCIAVAVVWAVF RNEDRWAWILQDILGJAFCINLIKTLKLPNFK SCVILLIGILLI.YDVFFYTIFTINGESBMYEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDILVFGL LIAYCRRFDVGTGSSTYYSVTVAYAIGMIL TFVVLGUMKKGQPALLVLVCTSLTIACQFV AWETVREMKFWERVTS CVILLIGULLYDFVFTTITNKGESBMYEL AGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDILVFGL LIAYCRRFDVGTGSSTYYSVTVAYAIGMIL TFVVLGUMKKGQPALLVLVCTSLTIACQFV AWETVREMKFWERVTS CVIPPSKGMTSAPALVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYNEAAAMNWRLFSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAMNWRLFSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLFSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLFSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSINFSAPLAVALTQT UNTPLLVYCAEYNEAAAAMNWRLPSKY QYPDSRGMFSAPSAAVCVAWASFQLTHG LKNCFI 624 1974 A 4622 164 668 VSCYTALQSIMMQPESANDPEPLCAVCQQAH SELENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTIFLVEKDFCFMBCTQVL QRCLEHHFQTSQAWGTHL*SQLLGRLQED CLSPSVHCHSEV CLS	1						<u> </u>
DPPKPSLSSQQAVYGHVHLGTGDPYTGGFPS	I	l	l		1	1	
NOTOFPPYASSGEPSPAQPISADIASRILIRKI KGPVAPQEWOGSLIGSPYHLGPGPRLRLVVN NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSMYSNIFG PRISILIFISWDGGDFCOSTEWLEGYLSVL HIKAVVYVSLDNAVLGDDKFHAKTSPLLTSIL IESVIKQVDSPNISGGILYEQVVFTNNPSWDI AEVIRPLPMDSSAYSFTAFVGVPAVEFSFMEN DDQNAYPFLHTKEDTYENLHVLGGRLPAVA QAVAQLAGQLLIRLSHDRILPLDFGRYGGDVV RAAEKLRQEIYSSEERDERLTRMYNYRIMRV EFYFLSQVVSPADSPFRHIPMGRGDHTLGALL DHIRLLRSNSSOTPGATSTGPQVESRFRRQLV ALLYWDACKGAANALSGDVWNIDINNP FOR STATE A 4610 793 334 ISRVDDFVGSGIANVILAVAITSIJPÄARLVRGF A 4610 793 334 ISRVDDFVGSGIANVILAVAITSIJPÄARLVRGF ALLYWDACKGAANALSGDVWNIDINNP EWGAMLARADMVAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG GSSIVVFFTMRIGTSIISAASLSFLGJGAQPTTP EWGAMLARADMVAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG LVYVMIAIFCIASAMSLYNCLAALHKIEPYGQ CTIACRGKMEVRLIFISGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLKTLKLPNFK SCVILIGILLLTVDFVFTITPTITKNGESIMVEL AAGPFGNNEKNDONLVATGOPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LAYCRRFDVQTGSSVTYVSVTVAVAIGMIL TFVVLGLMKKGQPALLYLVPCTLITACQFV AWETVREMKFVPENTPTITKNGESIMVEL AAGPFGNNEKNDONLVATAGOR A 4619 17 691 TLVSVVEFVRRADITREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSSPVSIPAVLTOT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMWRLFSKY QFDSRGMFISIVFSAPLLVNAMIVVMWW KTLNVMTDLKNAQERKEKKRRKED*GAA AAVSLRPSPPSAAPSAAVCVAWASFQLTHG LKNRCFI C624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQAULDPLD TPCGHTYCTLCLITNFLVEKDFCPCPMDRKPLVL QRCDLEHFQTSQAWGTHL*SQLLGRLQED CLSPVIHICSSE CVLSVFILPPTTLPSTLP C625 1975 A 4625 474 473 CVLSVFIPPLIFLISSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSSPSFFLPPPTTLPSTLP CFLSSSPLLPPILLLSSSSSPSFFLPPPTTLPSTLP CFLSSSPLLPPILLSSSSSSPSFFL		ł	1				
KGPVAPQEWQGSLGSPYHLGPGPRLRI.VVIANN NHRTSTPINNIFGCIESEDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSMVSNGFR PRRSILFISWDGGDFGSVGSTEWLEGYLSVIL HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL IESVLKQVDSPNISGQILYEQVVYTNIPSWDJ AEVIRPLPMDSSAVSFAPTAVGVPAVEFSFMED DDQAYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRILPJDFGRYGDVV LRHIGNLNEFSGDLKARGLTLQWVYSARGDY IRAAEKLRQEIVSSEERDERLTRMYNVRIMRVV EFFYELSQVYSPADSPRHIFMGROJPTIT GALL DHLRLLRSNSGTPGATSSTOFQESRFRQLA ALLTYNDACKGAANALSGDVWDINNIF ISSVDDFVGSGIANVILARHILPGT GSSIVVFFMRGITGSISASLSFLGLGAQFPTP EWGAMLNEARADMVIAPHVAVPPALAIFLTV LAFRILGDGLRDALDPKING 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNGLAALJHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVWAVF RNEDRWAWILQDILGIAFCINLIKTLRLPNFK SCVILLGLLLYDVFFVFTITFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGGPSAPHEKLPV VIRVPKLIJYFSVMSVCLMPVSLGFGDIVPGL LIAYCRFFDVGTGSSVTYYVSVTVAYAIGMIL TFVVLGLMKKGQPALVLYDFLITACQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIFAVLTOT DWTEPWLMGLATFHALCVLLTCLSRSSVRLQ IGHFLCLVILLVZCAFYNEAAAMNVRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KITLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPSAPSAAVCVAWASFQLTHG LKNRCFI LKNRCFI CG24 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPBEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPDRKPLVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLLVTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QRCKSSILVNKLINKLINTLTCPFRBETCTQVL QR	J	j]				
NHRTSTPNNIFGCIEGRSEPDHYVVIGAORDA WGPGAAKSAVGTAILLELVRTFSSMVSNGFR PRRSLIFISWDGGDFGSVGSTEWLEGYLSVL HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL IESVLKQVDSPNISGGTLYEQVYFTNIPSWD\ AEVIRPLPMDSSAYSFTAFVGVPAVEFSFME\ DDQAYPFLHTKEDTYENLHVLGGRLPAVA QAVAQLAGQLLIRLSHDRILPLDFGFRYGDVV LRHIGNINEFSGDLKARGLTLQWYSARGDY IRAAEKLRQEIYSSERDERLTRMYNYRIMRV EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQESRFRQDL ALLTWYDACKGAANALSGDVVNIDNNF FYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQESRFRQDL ALLTWDACKGAANALSGDVVNIDNNF 621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\ NTLVLKQOTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFALLGGLRDALDFRIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLAVAVVWAVF RNEDRWAWILQDLIGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITFITKNGESIMVEL AAGPFGNNEKNDGNLYLAFAGOPSAPHERLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYSVVTVAVAIGMIL TFVVLGUMKKGQPALLYLVPCTLITACQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIFAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVLVYCAESYINEAAAMNVRLJFSKY QYFDSRGMFISIVFSAPLLVAMIIVVMWWW KTLNVMTDLKNAQERRKEKKRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPPMDRKPLVL QRCKKSSILVNKLLINJCTCFFREHCTQVL QRCKSSILVNKLLINJCTCFFREHCTQVL QRCKSSILVNKLLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFFREHCTQVL QRCKSSILVNKLINJCTCFTREHCTQVL QRCKSSILVNKLINJCTCFTREHCTQVL QRCCTCHTATATATA							
WOGGAAKSAVGTAILLELVRTFSSMVSNCFR PRSLLFISWDGGDGSVSTEWLEGYLSVL HLKAVYYVSLDNAVLGDDKFHAKTSPLLTSL ESVLKQVDSPNHSGQTLYEQVVPTNPSSMV] AEVIRPLPMDSSAYSFTAFVGVPAVERSPME) DDQAYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRILPLDFGRYGDVV RHIGHLNEFSGDLKARGLTLQWYYSARGDY IRAAEKLRQEIVSSEERDERLTRMYNVRIMRV EFYYLSQVYSPADSPFHIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFOGSRFRRQL\ ALLYWDACKGAANALSGDVWNIDNNF GSSIVVFFMRIGTSIBAFARLVRG\ NTLVLKQQTFIESARSIGASDMYLLRHILPGT GSSIVVFFMRIGTSISASLSFLGLGAQPTP EWGAMLNEARADMYJAPHVAVFFALAIFLTV LAFNLLGDGLRDALDPKIKG AGOPTP EWGAMLNEARADMYJAPHVAVFFALAIFLTV LAFNLLGDGLRDALDPKIKG CTIACRGKNMEVRLIFLSGLCLAVVWAVF RNEDRWAWILQDILGIAFCLNLKTLKLPNFK SCVILLGLLLYDVFFVFITFFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGGPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIVPGL LIAYCRRFDVQTGSSTYTYVSVTVAYAIGMIL FFVVLGUMKKGPPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS GGCCESGLPHTMPSSRYSSSPPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSVRLQ IGHFLCLVLLVACAETVIEAAAAMNWRLPSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRSPSPPSAPSAAVCVAWASFQLTHG LKNRCF! CGC4 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENIFYSYPBEVDDDLICHICLQALDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRRPLVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLLVTCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPFREHCTQVL QHCKKSSILVNKLINKLINTLCPTPHCTQVL QHCKKSSILVNKLINKLINTLCPTPHCTQVL QHCKKSSILVNKLINKLINTLCPT	1	1					, -
PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL HLKAVVYVSLDNAVLGDKFHAKTSPLLTSL IESVLKQVDSPNHSGQTLYEQVVPTNRPSWD AEVIRPLPMDSSAYSFTAFVGVPAVEFSPME DDQAYPELHTREDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRLIPLDFGRYGDVV LRHIGNINEPSGDLKARGLTLQWVYSARGDY IRAAEKLRQEIYSSERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRIIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQESFRRQL) ALLYWDACKGAANALSGDVWNIDNNF 621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIGVAHSIPPAFARLVRG NTLVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPITVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKMEVELIFISGLCIAVVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPITKNGESIMVEL AAGPFGNNEKNDGNLVEATOGPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIVPGL LLAYCRFDVQTGSSYIYYVSVAVTVAYAIGMIL TFVVLGALMKKGQPALLYLVPCTLITACQFV AWETVREMKKFWERVTS 1745VLGALMKGQPALLYLLTCLSSRSYRLQ IGHFLCLVILVVCAGLTEDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEFWLMGLATFHALCULTCLSSRSYRLQ IGHFLCLVILLVYCAFURAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPSAAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDLICHICLQALLDPLD TPCGHTYCTLCLTIFILVEKDPCPMDRKPLVL QHCKKSSILVNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLVTCPTREHCTQVL QRCKKSSILVNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKKSSILVNKLLNKLLVTCPTREHCTQVL QRCKSSILVNKLLNKLLVTCPTREHCTQVL QRCLEHHFGTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV	1		ĺ	1			
HILKAVVYVSLDNAVLGDDKFHAKTSPLLTSL IESVLKQVDSPNHSGQTLYGVVFTNNFSWD AEVIRPLPMDSSAYSFTAFVGVPAVEFSPME DDQWYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLIRLSHDRLIPLDFGRYGDVV LRHIGNINEFSGDLKARGLTIQWYSARGDY RAAEKLRQEIYSSERDERLTRMYNVRIMRV EFYFLSQVYSPADSPFRHIPMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQ'ESRFRQL ALLITWDACKGAANALSGDVWNIDNNF GRIPVELSGAANALSGDVWNIDNINF GRIPVELSGAANALSGDVWNIDNINF GRIPVELSGAANALSGDWNIDNINF GRIPVELSGAANALSGDWNIDNINF GRIPVELSGAANALSGDWNIDNINF GRIPVELSGAANALSGDWNIDNINF GRIPVELSGAANALSGDATVLLRHILPGT GSSIVVFFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLLTLRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGASDMTVLTRHILPGT GSSIVVFTERISGAS	1						1
	ł	ł		}			
AEVIRPLPMDSSA YYSTTAPÖQVPA VERSFMEN DDQVAYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLIRLSHDRLLPLDFGRYGDVV LRHIGNINEFSGDLKARGLTLQWVYSARGDY IRAAEKLRQEIYSSERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLIRSNSSGTPGATSSTGFQUESRFRQU. ALLIVWDACKGANALSGDVWIDNINF 621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRGI NTLVLKQQTFIESARSIGAGSDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFTIFFTITKNGESMVEL AAGPFGNNEKNDGNLVEATQOPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYVSVITVAYAIGMIL TFVVLGUMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFREVENTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSFFVSIPAVLTQT DWTEPWLMGLATHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLPSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRREDFGGAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHICKKSSILVNKLINKLLVTCPFREHCTQVL QRCDLEHHFGTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSFSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP				1			
DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVV LRHIGNILPEDFGRYGDVT LRHILPEDFT EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLLRSNSSGYPGATSSTGFQESRFRRQL\ ALLITWDACKGAANALSGDVMNIDNNF ALLITWDACKGAANALSGDVMNIDNNF SIRVDDFVGGGIANVIIAVAIFSIPAFARLVRG\ NTLVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622	ļ		ŀ	1			IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
QAVAQLAGQLLIRLSHORLIPLDFGRYGDVV	,	į	l				
QAVAQLAGQLLIRLSHORLIPLDFGRYGDVV	İ	i			'		DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV	1	İ	i	1 !			
IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV							LRHIGNLNEFSGDLKARGLTLOWVYSARGDY
EFYPLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQUESRFRQL)			ļ	}			
621 1971 A 4610 793 334 ISRVDDFVGGSGIAVVIIAVAIFSIPAFARLVRG NTLVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTPT EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPPKK SCVILLGLLLLYDVFFVFTTPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIFYSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYVVSVITVAYAIGMIL TFVVLGLMKKGQPALLVLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFEVRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSPVSIPAVLTQT DWTEPWLMGLATTHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAFYNEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP				i l			-
ALLYWDACKGANALSGDVWNIDNNF				}			
1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\ NTLVLKQQTFIESARSIGASDMTVLLRHILPGT\ GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTP\ EWGAMLNEARADMVIAPHVAVFPALAIFLTV\ LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ\ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF\ RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK\ SCVILLGLLLLYDVFFVFITIFTITKNGESIMVEL\ AAGPFGNNEKNDGNLVEATGOPSAPHEKLPV\ VIRVPKLIYFSVMSVCLMFVSILGFGDIIVPGL\ LIAYCRRFDVQTGSSYIYYVSVVTVAYAIGMIL\ TFVVLGLMKKGQPALLYLVDCTLITA/CQFV\ AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF\ GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT\ DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ\ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY\ QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW\ KTLNVMTDLKNAQERKKEKRRKEKBCB-GAA\ AAWSLRSPSPPSAAPSAAVCVAWASFQLTHG\ LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH\ SLEENHFYSYPEEVDDDLICHICLQALLDPLD\ TPCGHTYCTLCLTINFLVEKDFCPMDRKPLVL\ QHCKKSSILVNKLLNKLLVTCPFREHCTQVL\ QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED\ CLSFGVHICSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	[ĺĺ			
MILVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVPPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFTTPFITTKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DVTEPPWLMGLATTHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPPREHCTQVL QRCDLEHHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPPREHCTQVL QRCDLEHHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPPREHCTQVL QRCDLEHHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPPREHCTQVL QRCDLEHHFYSYPEEVDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPPREHCTQVL QRCDLEHHFYSYPEEVDDLICHICLSCRIRGED CLSPGVHHCSEV	621	1971	Δ -	4610	703	224	
GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVJAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LYYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLYDVFFVFTITFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVTVAYAIGMIL TFVVLGLIMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP	1 02.	17/1	ΙΛ.	4010	133	334	
622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTTACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITFPTITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSVATVAYAIGMIL TFVVLGUMKKGQPALLYLVPCTLITA/CQFV AWETVERWKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP							
LAFNLLGDGLRDALDPKIKG)			
622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTTACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFTTPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTIGSSYTYYVSVVTVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP				}			
CTTACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP						-	LAFNLLGDGLRDALDPKIKG
RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATTHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLV\L QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
SCVILLGLLLLYDVFFVFITFFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLGUMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEFWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP							CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
SCVILLGLLLLYDVFFVFITFFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLGUMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEFWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP					ļ	ł	RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYIYYVSVVTVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATTHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							
VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSVVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATTHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							
LIAYCRRFDVQTGSSYIYYVSVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							
623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							
AWETVREMKKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP						ľ	
1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV GFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP					i		
GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP	623	1973		4610	17	691	
DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP		17.3	^	1017	• ′	971	
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							· · · · · · · · · · · · · · · · · · ·
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							· 1
KTLNVMTDLKNAQERRKEKKRRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP			1				
AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP			İ		į		
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP		}	ļ		j	ł	
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP	1	į		l		1	
SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLLPSTLP							
SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	624	1974	Α	4622	164	668	
625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	ļ	1		į	ŀ	1	SLEENHFYSYPEEVDDDLICHICLOALI DPI D
QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	1	į	- 1	ļ	i	ł	
QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV 625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP		į	-		į	l	
625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTTLPSTLP	1			1		ŀ	
625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP		ĺ	[1	ĺ	
	625	1075		4626	474	472	
PPLLIPSS*LSP	U2J	נועו	^	4043	4/4	4/3	
							PPLLIPSS*LSP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP ASASQVAGIAGTHH
627	1977	Α	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP ARAYL
628	1978	A	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF SYKSFAVIIFFVDNTRFFSFGF
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ EPALHRVGVYLDRGTGNVSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG FHDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG
632	1982	Α	4678	34	314	RSTHASGMISPSFOFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	Α	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFITYRITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYNVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH
635	1985	A	4709	42	341	YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	Α	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1]	sequence		nucleotide insertion
				Sequence		ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
	İ	Ì				MYFTTPSNHNAYQVDSVQST
637	1987	Α	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
			i			DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
ļ			}			HAG*AGLELLTSGDPPASASRSAGITGVSHHA
						RPRETRFL
638	1988	Α	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
ĺ		1				TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
						YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
İ			ŀ			LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
						WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITATICSLELPGSIDLPTSA
***	''''		'' ''	10,0	0,5	S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
	}	ļ				AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
•						TSNHVLYTQEGLRRGKEG
640	1990	Α	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
						WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS
						QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
						MARSRLTATSASQVQAILLPQPPGTTDSCSPS PDHEQQPLSWVLPPPQKDMNPREQQVALGP
						QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
					-,-	LOLAASPYFSPSWAECPOPVPAGTHATWCLA
1						RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
		1				FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
	1000	ļ.,	4550		100	QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	Α	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
						TWWFGVKFAAGGLGTFHALLNTAVHVVMY
		1				SYYGLSALGPAYQKYLWWKKYLTSLQLVQF
						VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
		}				FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
		l	·			QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
		,				MVYFVGENNGDSSHNPVLAATGVGLDVAQS
						WEKAIRQALMPVTPQASVCTSPGQGKDHSK Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
*		^^	.500			LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
						AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
						SYKDIWGWPCLCGVLHAYIPLLV
645	1995	Α	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV
j i]				AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
						PLLAGLVAADAVASLLIVGAVFLCARPRRSP
646	1996	A	4817	47	1033	AQEDGKVYINMPGRG LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
070	1770	^	701/	71	1023	LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
						HKQAVQCLKGPGQVARLVLERRVPRSTQC
						PSANDSMGDERTAVSLVTALPGRPSSCVSVT
						DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
						KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
						WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
						YYPAAVEVLHILLRGAPQEVTLLLCRPPPGAL
						PELEQEWQTPELSADKEFTRATCTDSCTSPIL GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
						EGTMGAKTERDLGPVP
647	1997	A	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD
لــــنـــا						TELESTICAL CONTROL OF

NO: of No: of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation				,		li .	
Solid		l .	1100	i		1	
Sequence							
Before	•			1	1		, , , , , , , , , , , , , , , , , , , ,
amino acid residue of sequence colored residue of sequence colored residue of sequence colored residue of sequence colored residue of sequence colored residue of sequence colored residue colored	1 1	uence					
Peptide Petide Peptide Peptide Petide Peptide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Petide Pe	uence		}	914	, –		
Peptide Pept			1				
]					sequence	
					peptide		
			ŀ		sequence		nucleotide insertion
							IVMPTYDLTDSVLETMGRVSLDMMSVQANT
HPELIDAAFTNFFFFKHDENLYGPIVKHIS FFFKHKYQNIDGTVAAYALPLJUGDSV QDSIYYFHFYNELQPWKHYIPVKSNLSD LKWAKDHDEEAKKIAKAGQEFARNNIL DIFCYYFQTFPRNMPIYK SEEEKSGLGAKIPVIIITGYIQAGKTITLI LTECHSKRVAVILNEFGEGSALEKSLT GELYEFWLELRNGCLCCSYKDNGLRAIE QKKGKPDYILLETTGLAPPGAVASMFW ELGDIYLDGITTIDSKYGLLETGGAPGAVASMFW ELGDIYLDGITTIDSKYGLLETGGAPGAVASMFW ELGDIYLDGITTIDSKYGLLETGGAPGAVASMFW ELGDIYLDGITTIDSKYGLLETGGAPGAVASMFW ELGDIYLDGITTITSKYGLGETORSTVALDLIAI SGISLQKKLQHVPGTOPHLDGSIVTITFDI NAKFELLNGKLCSYKDNGLRAIE QKKGKPDYILLETTGLAPPGAVASMFW ELGDIYLDGITTITSKYGLGAPGAVASMFW ELGDIYLDGITTITSKYGLGAPGAVASMFW ELGDIYLDGITTITSKYGLGAPGAVASMFW ELGDIYLDGITTITSKYGLGAPGAVASMFW ELGDIYLDGILETQRSRVDLUBAI SGISLQKKLQHVPGTOPHLDGSIVTITFDI NAKFELLNGKLGVWQVWAHWOPHLDGISTTTTDI NAKFELLNGKLGVWQVWAHWOPHLDGISTTTTDI NAKFELLPGAVAYAHWAHWOPHLCVITTAKINGLAPTICATA STRILLPROSECTION TETTER]		l	l	}	ł	GPPWESKNSTAVWRGRDSRKERLELVKLSRK
FINHKYQNIDGTVAAYRLPYLLVGUS			İ	ĺ		<u> </u>	
				i			II.
					ł	ì	
				1			1 7
	1 1			[[
SEEEKSGLGAKPYTIITGYLGAGKTTLI LTEQHSKRVAVILNEFGGESALEKSLAVS GELYEEWLELRNGCLCCSVKDNGLRAIE GEKYEEWLELRNGCLCCSVKDNGLRAIE QKKGKPDYILLETTGLADPGAVASMFW ELGSDIVLDGIITIVDSKYGHLADEKPK NEATRQVALADAILINKTDLVPEEDVKK TIRSINGLGQILETGRSRVDLSSNVLDHAI SGISLQKKLQHVPGTQPHLDQSIVTITIFDI NAKEEHLNMFIQNLLWEKNVRNKDNHC RILKGLVSIKDKSQQVIVQGVHELYDLEE SWKDDTERTINRLVLLGRNLDKDILKQLF VIETEKQWTHPKEDQVIVGVHELYDLEE SWKDDTERTINRLVLLGRNLDKDILKQLF VIETEKQWTHPKEDQVIVGVHELYDLEE SWKDDTERTINRLVLLGRNLDKDILKQLF VIETEKQWTHPKEDQVIVGVHELYDLEE SWKDDTERTINRLVLLGRNLDKDILKQLF VIETEKQWTHPKEDQVIVGVHELYDLEG GQAGLELRTSGDPPASASQSAGITGVSHII TSMPLLPPQRLCVYI FFLRRSFAVAQAQVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFLRRSFAVAQAQVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDYRIIAPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDWRIIAPPCPS*FLVF** FFMLRSFAVAQAGVQWCDLGSPQPLP K*PSCLSLPSSWDWRSKSSDPVLP RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPFTQUSPTSGIC SP*ILPCASNRLAFGGLIFPCAPLVPYPAPP APSCAPPRAHTHSRTHPALVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC K*PGDSWRVAACISGRCCSRGRRGSGFR SS*ILPCASNRLAFGGLIFPCAPLVPYPAPP PAFSCAPPRAHTHSRTHPALVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC K*PGDSWRVAACISGRCCSRGRRGSGFR SS*ILPCASNRLAFGGLIFPCAPLVPYPAPP PAFSCAPPRAHTHSRTHPALVPKPSSI GGSPIPSRASSPSCSWAQVPGVALARCAC K*PGDSWRVAACISGRCCSRGRRGSGFR SS*ILPCASNRLAFGGLIFPCAPLVPYPAPP APSCAPPRAHTHSRTPAPLVPKPSSI GGSPIPSRAGSFRGWGSTPGGSGFR SS*ILPCASNRLAFGGLIFPCAPLVPYPAPP TSCASNRLAFGGLIFPCAPLVPYPAPP TSCASNRLAFGGLIFPCAPLVPYPAPP TSCASNRLAFGGLIFPCAPLVPYPAPP TSCASNRLAFGGLIFPCAPLVPT TSCASNRLAFGGLIFPCAPLVPT TSCASNRLAFGGLIFPCAPLVPT TSCASNRLAFGGLIFPCAPLVPT TSCASNRLAFGGLIFPCAPLVPT TSCASNRLAFGGLIFPC	i-,,,	1000	<u> </u>	4967	2020	027	
	040	1998	l A	4007	2030	837	,
GELÝZEWLEIRNGCLCCSVKDNGLRAIR GRELÝZEWLEIRNGCLCCSVKDNGLRAIR GKKGKFDYILLETTGLADPGAVASMFW ELGSDIYLDGIITTUPSKYGLKHLAEEKPI NEATRQVALADAILINKTDLVPEDVKK TIRSINGEGQILETGRSRVDLSNVLDLHAI SGISLQKKLQHVPGTQPHLDQSIVTITFDI NAKEEHLJMFIGPILLWENVRNKDYDLHAI SGISLQKKLQHVPGTQPHLDQSIVTITFDI NAKEEHLJMFIGPILLWENVRNKDJLDLKQLF RILKGLVSIKDKSQQVIVQGVHELYDLEE SWKDDTERTNRLVLLGRNLDDILKQLF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETEKQWTHIFKEDQVCT OF VIETKQUT AND AND AND AND AND AND AND AND AND AND	1		1	ļ			•
QKKGKPDVILLETTGLADPGAVASMFWP ELGSDIYLDGIITVDSKYGLKHLAEKEKP NEATRQVALADAILNKTDLVPEEDVKK TIRSINGLGQILETQRSRVDLSNVLDLHAA SGISLQKKLQHVPGTQPHLDQSIVTITFD NAKEEHLJMFIQNILLWEKNVRNKDINGL IRLKGLVSIKDKSQQVIVQGVHELYDLEE SWKDDTERTNRLVLLGRRIDDKDLKQLF VTETEKQWTTHFKEDQVCT OF SCLLLPKLGQVWAQYWAHWQPPLPC SWKDDTERTNRLVLLGRRIDDKDLKQLF VTETEKQWTTHFKEDQVCT OF SCLLLPKLIGQVWAQYWAHWQPPLPC SWKDDTERTNRLVLLGRRIDDKDLKQLF VTETEKQWTTHFKEDQVCT OGAGLELRTSGDPPASASQSAGITGVSHILTSHPPFFLVF4* TSMPLLPFQRLCVYI 650	l :			l			1 .
			!				j
NEATRQVALADAILNIKTDLVPEDVKK TIRSINGLGQILETQRSRVDLSNVLDLHAI SGISLQKKLQHVPGTQPHLDQSIVTITFD NAKEEHLMMFIQNILWEKNVRNDSNHC IRLKGLVSIKDKSQQVIVQGVHELYDLEE SWKDDTERTNRLVLLGRNLDKDILKQLF VTETEKQWTTHFKEDQVCT				[}	QKKGKFDYILLETTGLADPGAVASMFWVDA
TIRSINGLGGILETQRSRVDLSHVLDLHAL SGISLQKKLQHVPGTQPHLDQSIVTITFD NAKEEHLNMFIQNILWEKNVKNKDHC RILKGLVSIKDKSQQVIVQGVHELYDLEE SWKDDTERTINILVLLGRID.BDDILKQLF VTETEKQWTTHFKEDQVCT SVENDTERTINILVLLGRID.BDDILKQLF VTETEKQWTTHFKEDQVCT SVESLLIPKLGVQWAQYWAHWQPPLFF SCLSLRSSWD*KAPPHPAFVFLVEMGF GQAGLELRTSGDPPASASQSAGITGVSHI TSMPLIPFQRLCVYI TSMPLIPFQRLCVYI TSMPLIPFQRLCVYI TSMPLIPFQRLCVYI FFELRSFARVAQAGVQWCDLGSPQPLPF K*FSCLSLPSSWDVRIIAPPPCPS*FLYF*** FTMLARLVLNS*PHDLTPSSPQSAEIKGV CPASFYLFLKYYLEAKFCA*GECAFSAGV GYKRGHKSCLLINCVVQI GYKRGHKSCLLINCVVQI AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO AALKLPILVSKILDIPGLEVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO PHPVAVTTERAF PHMEPKASCPAAAPLMERKFHVLVGVTO							ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI
SGISLOKKLÖHVPGTOPHLDQSIVTITFDV NAKEEHLNMFIQNLLWEKNVRNKDNHC RILKGLVSIKDKSQVIVQGVHELYDLEE SWKDDTERTNRLVLLGRNLDKDILKQLF VTETEKQWTTHFKEDQVCT SWKDDTERTNRLVLLGRNLDKDILKQLF VTETEKQWTTHFKEDQVCT FSCLSLRSSWD*KCAPPHPAFVFLVEMGI GQAGLELRTSGDPASASQSAGITGVSHI TSMPLLPFQRLCVYI FSCLSLRSSWD*KCAPPHPAFVFLVEMGI GQAGLELRTSGDPASASQSAGITGVSHI TSMPLLPFQRLCVYI FFELRRSFAFVAQAGVQWCDLGSPQPLP FFELSLSPSSWDYRHAPPPCPS*FLYF** FTMLARLVLNS*PHDLPTSPSQSAEIKGV CPASFYLFLKYYLEAKFCA*GECAFSAGI GYKRGHKSCLLINCVVQI A 4898 1701 771 DAWGPETRLARILNPDSFIEPRPGRLPELE PHMEPKASCPAAAPLMERKFHVLVGVT AALKLPLLVSKLIPDGLEVAVYTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLF RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLPCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGF EVGTIVDKVKBULPGPSGGFSPKGSFPFTPGFPGFSGG EVGTIVDKVKBULPGPSGGFSPFFFFFF FFFFFFFFFFFFFFFFFFFFFFFFFFFF			1	1			NEATRQVALADAILINKTDLVPEEDVKKLRT
NAKEEHLNMFIQNIL.WEKNVRNKDNHC IRLKGLVSIKDK SQQVIVQGVHELYDLEE SWKDDTERTINK VLLGRNLDKDILKQLE VTETEKQWTTHFKEDQVCT G49				1			TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL
NAKEEHLNMFIQNIL.WEKNVRNKDNHC IRLKGLVSIKDK SQQVIVQGVHELYDLEE SWKDDTERTINK VLLGRNLDKDILKQLE VTETEKQWTTHFKEDQVCT G49]			l			SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG
SWKDDTERTNRL VILGRILDKDILKQLE				1			NAKEEHLNMFIONLLWEKNVRNKDNHCMEV
SWKDDTERTNRL VILGRILDKDILKQLE				l			IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
		•		1			
1999	,		[)			
FSCLSLRSSWD*KCAPPHPAFVFLVEMOR	649	1999	A	4873	226	189	
GQAGLELRTSGDPPASASQSAGITGVSHI TSMPLLPFQRLCYYI		•,,,,	1.	10/3	220	107	
TSMPLLPFQRLCVYI				ŀ			4
A							1 7
K*FSCLSLPSSWDYRIIAPPPCPS*FLYF** FTMLARLVLNS*PHDLPTSPSQSAEIKGV. CPASFYLFLKYYLEAKFCA*GECAPSAGV. GPKRGHKSCLLINCVVQI 651 2001 A 4898 1701 771 DAWGPETRLARILNPDSFIEPRPGRLPELE PHMEPKASCPAAAPLMERKFHVLVGVTG AALKLPLLVSKLLDIPGLEVAVYTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPIT VQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWQAKSVKMDVGKIGGYPHLLNG LSLPRQQACSRLNWTEGPGLSFFOPGEAA LSLPRQQACSRLNWTEGPGLSFFOPGEAA SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSIG GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFFFDGVSLCHPGWNAVARSWLTATS/VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTLAILVINS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RFGIPGRFRRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTGRSQGR DPERDPGCRVSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	650	2000		4074	3	427	
FTMLARLVLNS*PHDLPTSPSQSAEIKGV:	030	2000	A	46/4	2	437	
CPASFYLFLKYYLEAKFCA*GECAPSAGY GYRRGHKSCLLINCVVQI 651 2001 A 4898 1701 771 DAWGPETRLARILNPDSFIEPRPGRIPELE PHMEPKASCPAAAPLMERKFHVLVGVTY AALKLPLLVSKLLDIPGLEVAVVTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSFWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFF1*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHAALVLNS*PQVICPPWPPKVLTLC 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEGLESLAT IPAVGLGALGVIPPVRVPQRPTTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							,
GYKRGHKSCLLINCVVQI		1		ł			-
651 2001 A 4898 1701 771 DAWGPETRLARILNPDSFIEPRPGRLPELE PHMEPKASCPAAAPLMERKFHVLVGVTG AALKLPLLVSKLLDIPGLEVAVVTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEA/ EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEA/ FRGPSPRAFGFSPWRPPGTMQEPSSGF SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSF GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTHAILVLNS*PQVICPPWPKVLTLQ WGFTHAILVLNS*PQVICPPWPKVLTLQ WGFTHAILVLNS*PQVICPPWPKVLTLQ GSFGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI				ŀ			
PHMEPKASCPAAAPLMERKFHVLVGVTG AALKLPLLVSKLLDIPGLEVAVVTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPITI VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRQQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHIPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YTT WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEFGLESLAT IPAVGLGALGVIPVRVPQRPTTQRSQGR DPERDPGCRVQVSRGPRRGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
AALKLPLLVSKLLDIPGLEVAVVTTERAK SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMGPSSGE GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	651	2001	Α	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR
SPQDIPVTLYSDADEWEMWKSRSDPVLH RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLLFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA LSLPRGQACSRLNWTEGPGLSFFQPGEAA LSLPRGQACSRLNWTEGPGLSFFQPGEAA SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPEGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGRG DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							PHMEPKASCPAAAPLMERKFHVLVGVTGSV
RRWADLLLVAPLDANTLGKVASGICDNI VMRAWDRSKPLIFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSPWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
VMRAWDRSKPLLFCPAMNTAMWEHPIT VDQLKAFGYVEIPCVAKKLVCGDEGLGA EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSE GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL
VDQLKAFGYVEIPCVAKKLVCGDEGLGAEVGTIVDKVKEVLFQHSGFQQS*PGISVMLYSEWVQAKSVKMDVGKIGGYPHLINGLSPRGQACSRLNWTEGPGLSFFQPGEAAFSCAPRAGFSPWRPPGTMQEPSSGEAFSPLPCASNRLAFGGLIFPCAPLVPYPAPFPAFSCAPRPRAHTHSRTHPSAPLVPKPSSGEAFSPRGSGPRQSFPRASPSCSWAQVPGVALARCAGEAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG							RRWADLLLVAPLDANTLGKVASGICDNLLTC
EVGTIVDKVKEVLFQHSGFQQS*PGISVM LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSGE GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PA*FSCAPRPRAHTHSRTHPSAPLVPKPPSF GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA
LYSEWVQAKSVKMDVGKIGGYPHLLNG LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PA*FSCAPRPRAHTHSRTHPSAPLVPKPPSF GQSPIPSRASSPSCSWAQVPGVALARCAG KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP
LSLPRGQACSRLNWTEGPGLSFFQPGEAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSE GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFF1*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTH_AILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							LYSEWVOAKSVKMDVGKIGGYPHLLNGGPA
652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGE SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSE GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTH_AILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							LSLPRGOACSRLNWTEGPGLSFFOPGEAAA
SP*LPCASNRLAFGGLIFPCAPLVPYPAPF PAFSCAPRPRAHTHSRTHPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTH_AILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFFRSWFCQLP*EPEGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	652	2002	A	4927	1	611	
PAFSCAPRPRAHTHSRTHPSAPLVPKPSSI GQSPIPSRASSPSCSWAQVPGVALARCAC KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI					-		SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
GQSPIPSRASSPSCSWAQVPGVALARCAGE KPGDSWRVAACISGRCCSRGRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPFGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
KPGDSWRVAACISGRCCSRGRRRGSGPR QSFRGAWGPSFWGSWKSQRELSAGGAQ LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
G53 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSA VQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTH_AILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGR DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSAVQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTH_AILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPGRVGRAGGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPRFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQVSRGPFGEQKTPGLQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQGALGVIPPVRVPQRPTQRSQGROUPERDPGCRVQGALGVIPPVRVPQRPTQRSQGALGVIPPVRVPQRPTQRSQGALGVIPPVRVPQALGVIPPV	1						
653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSAVQAVSCFRLPSSWDYRHATMPG*FF*YF*WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGRO DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
VQAVSCFRLPSSWDYRHATMPG*FF*YF* WGFTHLAILVLNS*PQVICPPWPPKVLTLQ 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGRI DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	452	2002	Α	4005	2	202	
654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGRUPPCLTHLAAASCVVVWCGRWKRDSAECHSCSAVSQQEDRCRSSSCS A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	000	2003	A.	4900	2	283	
654 2004 A 4968 3 . 437 RPGIPGRRFRSWFCQLP*EPEPGLESLAT IPAVGLGALGVIPPVRVPQRPPTQRSQGRUPPCLTHLAAASCVVVWCGRWKRDSAECHSCSAVSQQEDRCRSSSCS A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
IPAVGLGALGVIPPVRVPQRPPTQRSQGRI DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI				10.65			
DPERDPGCRVQVSRGPRFGEQKTPGLQG PPCLTHLAAASCVVVWCGRWKRDSAEC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	654	2004	A	4968	3 .	437	
PPCLTHLAAASCVVVWCGRWKRDSAÈC HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							
HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	ļ						DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP
655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI							PPCLTHLAAASCVVVWCGRWKRDSAECQCD
655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVI	l						HSCSAVSQQEDRCRSSSCS
	655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN
, , , , , , , , , , , , , , , , , , ,			ļ				TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
	656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI
REVHIKTMR*HFLPIRLEKNKNNIKD							
	657	2007	R	รถกร	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES
557 2507 D 5000 127 403 MAGINATASOD I DSS WELKY F VOEED FE		2007	٠	2000	147	-05	MANOWIN I MOOD I INGO WELL AL A GEENLEVE?

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
•		1	1	residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
ļ				peptide	scquence	/=possible nucleotide deletion, \=possible
	ł		l '	sequence	}	nucleotide insertion
<u> </u>	 		 	sequence		VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
						AIWWEQKRQWLLQTHWTLDKYGILADARLF
İ		ĺ		İ		FGPQHRPVILRLPNRRALRLX*
658	2008	Α	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
	1					KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
						HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	Α	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
	ŀ	ŀ				T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
]		}	į		j	MAACWAVHVKTHMRPGLAVLPRLVLNSWS
	0010	l	5000		210	*AIILLWPPKALGLQA
660	2010	Α	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
		1			1	HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST AGSCPROKKTTPGPTVLCVCSFWIYORGEPH
1	Ì	(HRTGARWNH
661	2011	A	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
001	2011	^	3030	732	421	LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
						LELLGSSHPPTSASQSARITGVSHRAWPLK*F
		}				NLNOYOTLTMN
662	2012	Α	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
						EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
	ĺ					LHEACLGDHVACARTLLEAGANVNAITIDGV
			i			TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
						SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
İ						TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
(/2	2012		5066	051		WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	Α	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
						GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
}			}			WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	A	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
			***		•	QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
						MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
		٠				TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
						YMAICNPLRYSVLMGHGVCMGLMAAAWAC
						GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	Α	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
						PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
						HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
						HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
						KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
				İ		ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
	,					PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
						VGPKAQLSGLQLQPCLYKRR
667	2017	Α	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
				Į	_	PCVY*RR
668	2018	Α	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
						RALPTTFADIENLKYLLFTRDASQPFYLGHTV
						IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
						TCADQSVIWKLSEDKQLAICLKYAGVHAENA
						EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
						CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
660	2010		610:		320	GHYFNDTLVFLPPVGSEND
669	2019	Α	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
				ļ	.	ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP
				}		SPILPVLLPELGPL
670	2020	A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
						Z Gradell Gradital Golf ADI GRA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
						VGGTDQQYVSNDSGIYVSRIKENGAAALDGR LQEGDKILSVNGQDLKNLLHQDAVDLFRNA GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI FMVLVPVFALTMVAAWAFMRYRQQL
671	2021	A	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF VLLLLLISLLCLYWKARKLSTLRSNTRKEKA LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA NHFVEVT
674	2024	Ā	5153	3	2953	LTEDOPFDILQKSLQEANITEQTLAEEAYLDA SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG QTLQPIGVTHVPVGASFASNTVGVQHGFMQH VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS MMTINNLDGSQIILKGSGQQAPSNVSGGLLV HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM GQQNTYNVNNLGIQQHHVQQGISFASASSPQ GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ TFAASGSPVIANHASPQLVGGQMPLQQASPT VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE MVMIDRMFNQEERASLSRDKRLALVDPEGFQ ADFCCSFKLDKAAHETQFGRSDQHGSKASSS LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS RSDHGTESKLSSILADSHLEMTCNNSFQDKSL RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET TFKNILELKKAGRQPQSDPTVSGSVELDFPNF SPMASQENCLEKFIPDHSEGVVETDSILEAAV NSILEC
675	2025	A .	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR LYDEVQEVVYFPAVVHDNLGERLKCTYIEID QVPETYAVVLSRPAWLWGAEMGANEHGVCI GNEAVWGREEVCDEEALLGMDLVRLGLERA DTAEKALNVIVDLLEKYGQGGNCTEGRMVF SYHNSFLIADRNEAWILETAGKYWAAEKVQE GVRNISNQLSITTKIAREHPDMRNYAKRKGW WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE GYKLLNKHKGNITFETMMEILRDKPSGINME GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
676	2026	A	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG FTLLARMVSIS*PHDPPASASQSAGITGVSHRA RPT
677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR LNKRSFFMISPTDQQVHCWAWLKKIIMPKDS NLLLEDVTWKYTALNLIGPRAVDVLSELSYA PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	A	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ GRIAKMPVKWIAIESLADRVYTSKSDVWAFG VTMWEIATRGMTPYPGVQNHEMYDYLLHG HRLKQPEDCLDELCKI**SPQSP
679	2029	A	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS
680	2030	A	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK OSESAI
684	2034	A	5220	1	194	NLMKEMONLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS
685	2035	Α .	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT
688	2038	Α	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

NO; of No; of nucle unice unice with the provided sequence unice u	CCOID	LOFOID	1 3 2	1.000	T		
nucleotide sequence contide sequence peptide sequence peptide sequence peptide sequence continue sequence continue sequence peptide sequence peptide sequence continue sequenc	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
DISSN Dodation D	•		nod	1 .			
sequence sequence 09/496	1	1	1		B		F=Pnenylalanine, G=Glycine, H=Histidine,
ueince 914 ng to first amino acid residue of peptide peptide peptide sequence Pept				1			M=Methionine N=Amornaine D-Deline
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequences sequence sequences sequ							O=Glutamina D=Arginina S=So-ina
residue of peptide sequence Peptide sequence Peptide peptide sequence Peptide pertide pertide sequence Peptide pertide		ì	1				T=Threonine V=Voline W=Trustonhon
Popular Popu	1	ł	ł	1			Y=Tyrosine Y=Unknown *=Ston coden
	ļ		J	İ] sequence	/=nossible nucleatide deletion \=nossible
				•			
PKSLAPSTPPNSSHSPCPSILPEPEVHUSE FLASRAEFIL_QEMYGLRESNSPPDVGQLKM NLVDRIKRLGIARVVKNPGAQENGROGUSE				\top	-		
FLASRFAETFLQEMYGLRPSRNPPDVOGQLEANN NLVDPIARRICIDARVVKNPGAQPGIGCQEAE GPQRPDSAVYLNSGSSLLGGLRRNQSLPVIM GSFAAPVCTSSPKMQVLKED GPQRPDSAVYLNSGSSLLGGLRRNQSLPVIM GSFAAPVCTSSPKMQVLKED LSLFGSRALGRSGARAMAKAKVQARKKV GAPAGARGGGAKANSNPFEVKVRQKKEQTLL RKTRHDVGLPGVSRARALRKRTQTILKEYK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RAKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEEKMMKRK RDKSNVFRDKRGEVNSNMSPEELALAKEGGERLKKKEKKPPD V DMMVRELGF EMKAQPSNRMKTEAELAKEGGERLKANEGGGENEESDSPSHDLDLES RDKSNEEGESDSGGDTEESDSPSHDLDLES RDKSNEEGESDSGGDTEESDSPSHDLDLES RDKSNEEGESDSGGDTEESDSPSHDLDLES RDKSNEEGESDSGGDTEESDSPSHDLDLES RVSSEEEMKPAKEGRGTPGKGLISGKERAG KATKDELPYTFAAPESYEELASLLGRSMEE LLVVERIGKCHHPSLAEGMKAKLEKLEFFL VYGGLATDDPFLVIDKLVYLICOMF ESASDAIKFVLRDAMHEEMEMETIKGRAAU RUKVERLEY	ł		1	1	l	!	PKSLAIPSTPPNSPSHSPCPSPI PEEPRVHI SENI
NLVDRIKRLGIARVYKNYFGAQENGROGE GPQKPDSAVYLNGSSLLIGGLERNOSLPVIM GSFAPVCTSSFKMGVLKED SELEGRALIZESGRALMSKRKYGARRKAS GPKPDSAVYLNSGSSLLIGGLERNOSLPVIM GSFAPVCTSSFKMGVLKED GSFAPVCTSSFKMGVLKED SELEGRALIZESGRALMSKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS GAPAGARGGRAMAKKKYGARRKAS LEQQRHHEKKSIYNLINDEELTHYGQSLAMIK RDKSNYFROKRRGEVNSKMSPESKMKKRA LEQQRHHEKKSIYNLINDEELTHYGQSLAMIK RDKSNYFROKRRGEVNSKMSPESKMKRRA LEQQRHHEKKSIYNLINDEELTHYGQGSGC GLIHKKYQGGGEREKKSKKELIELLIAKGGG GLIHKKYQGGGEREKKSKKELIELLIAKGGG GLIHKKYQGGGEREKKSKKELIELLIAKGGG GLIHKKYQGGGEREKKSKKELIELLIAKGGG GLIHKKYQGGGEREKKSKKELIELLIAKGGG GLIHKKYQGGGEREKKKSKKELIELLIAKGGG GLIHKKYQGGGERIAKKEKKPKKK GEKKRQOSTINDEKSKEPPDAYDMMYRELLIAK GEMAPAGE GLINDERNIK KEKPKPDAYDMMYRELLIAK GEMAPAGE GLIPVTPAAPESYELISLLIARSMED LIVVERIGKCHPSJAGMKAKLEGGFLL VYGDLATIDDEPDITVIDKLIVVHLYHLCQMFF ESASDAIKFYLRDAMHEMEEMIETKGRAAL GAPTITLIAL GULTIKCTILIAL GULTIK			İ				FLASRPAETFI OEMYGL RPSRNPPDVGOLKM
GPQKPDSAVYLNSGSSLLGGLRRNGSLPVINK GSFAPVCTSSPKMQVLKED			ļ	•	ŀ		NLVDRLKRI.GIARVVKNPGAOFNGRCOFAFI
SSFAAPVCTSSPKMOVIKED			i	1	İ	l	GPOKPDSAVYLNSGSSLLGGLRRNOSI PVIM
\$2039 A \$254 2 \$2621		1	İ		l		GSFAAPVCTSSPKMGVLKED
GAPAGARGGPAKANNNFFEWVMRQKFQILL RKTRIDVOELPOVSBRARLARRITOTILLREYSIL RKRTRIDVOELPOVSBRARLARRITOTILLREYSIL RKRTRIDVOELPOVSBRARLARRITOTILLREYSIL RKRNEDVSDSDADAEDRGTI.SGELTAAHFGGGG GILLHKKTQGEGEREKPKSRKELIEELLAKSK QEKKRRQAQREDALELTEKJODWKEIGTIL SHKTIPKSENRDKKEKPKPDAYDMMYRELGF EMKAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILRILGAE ENKRAQPSNRMKTEAELAKEGHILAGSMEE LIVVERQKCNIHPSI.ABGNRKAKLEKILGGBLE YVGOLATIDPDITVIDKLVYHLYHLCQMFF ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLUVIYIKITGILLFYTSTSPWTYTALVCL SQLLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVALS QRLITKCPII.SLQDVVKGI.FVCCLFILEYVI.FV GELQSTI.TIEMESQKQI.CRP.TCEKSKP.FVLK. FTRIL.VVLI.EGRQCGSSCTI.VIPPT ALGRAMAGE.LAVINGAE.GRAM. SRLRAPTSTEAHHIR.SCLAVG.GRAVHLLKKVI.KVFKK REFKGAVREIKZDNOFI.ARMQ.SESILERVAL FTRIL.VVLI.EGRRQCGSSCTI.VITHORITAA SRLRAPTSTEAHHIR.SCLAVG.GRAVALIKKVFKK REFKGAVREIKZDNOFI.ARMQ.SESILERDAE ENKRINGATA.HAURA SRLRAPTSTEAHHIR.SCLAVG.GRAVALIKKVFKK REFKGAVREIKZDNOFI.ARMQ.SESILERDAE ENKRINGATA.HAURA KRYNCYLETPII.VVLI.FGTPFIGWKIDGPAAAL EVLSSFFFFFLKSSYKFQNIV 691 2041 A 5270 3 158 NCHTTCTANWVILPGTPFPGWKIDGPAAAL EVLSSFFFFFLKSSYKFQNIV FTRILARRATICHTAAHTALTHALTHALTHALTHAURALTHATAATAATAATAATAATAATAATAATAATAATAATAAT	689	2039	A	5254	2	2621	
RKTRIBVGLPGVSRRALKRKTOTILLEYKI RDKSNVFRDKRPGFYSNMSPEEKMMKKRRA LEQQRHHEKKSIYNINEDEEL THYGQSLADII KHNDIVDSDSDAEDRGTLSGELTAAHFGGGC GLHKKTQBGGEREKPKSRKELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIGTLI SHKTYKSEPRIKKERKPFDAYDMMVRELGF EMKAQPSNRMKTEAELAKEEGEHLRKLEAE RIRRMLGKDEDENVKKFKHMSADDLNDGF LLDKDDRRILSYKDGKMNVEEDVQEEGSKEA SDPESNEEEGOBSGGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQPTPGKGLISGKERAG KATRDELPYTTAAPESYEELRSLLLGRSMEEC LLVVERIQKCNHPSLAEGNKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVVHLVHLCOMFF ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLLFPTSDFWIPVVTPALVCL SQLLTKCPHISLQDVVKGLVECLFEYVALS QRFIPELINFLLGILYIATPNKASQGSTLVIPFF ALGKNISELLVVSAREDVATWQQSSLSIRWA SRRAPTSTEAHHRISCLAVGLALLKRCVLM YGSLPSFHAIMCPLRALITDHLADCSHPQELQ ELCQSTLTEMESQQGLCEPTCSKSPVPHLKL FTPRLVKVLEFGRKQGSSKEEGERKELHKHK REFKGANREIRSQGLGERLTSGSDPTW ASQSAGITGVSHCSWAVIVLSTLHAWRNVL FKRTFPLKSSSFLSYDKEIPPILIVLKFYLVTLT FFYL 691 2041 A 5270 3 158 NCHTHCTANWVHLPGTPFGWKIDCPAAAL FKRYKVOLFNSLATOGGEWKALKRKKKK KERFKGANREIRSGNGLCAPTUSTLHAWRNVL FKRTFPLKSSSFLSYDKEIPPILIVLKFYLVTLT FSYK 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLSTFFFFLKFSKYPONU KRITPLAKALLRVACVCRLWRECVRVLR THRSVTWISAGLAEGHLEGHCLVRVVAEEL ENVRLIPHTVLYMADSETFISLEECRGMKRAR KITSMETALALEKLPRKQCVLGIVTPGIVVT PMGSGSNRPQEEIGGSGFALLFPGEGIGPF FHIKKDPKINT.LERRIQLTEGULDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDANIILAGG QVSNLSSLTSEKNPLDIDASGVVGLSFSGIRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMACVGRGFQYYRAKGNVEADAFR KFFSVPLFGFGRGGIGGEORIVTGNFILKKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EERKERFGFGRUIVWYGFIQELDCNPERCRILK ACPPTINVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRRYFYKGKKKKPKELLPPE KTAMMRENLKKYKKKKRKKCHRKERLYPL FRYNTENLERGLTRYFYKGKKKKPKCHLERERLYPE FRYNTERFORGEGICTIVTFYFRGKKKKHTERLYPE FRYNTERFORGEGICTIVTFYFRGKKKCHRERLYPL FRYNTENLERGERGELLTRYFYKGKKKKFRENLYPL FRYNTENLERGYNTERGELLK FKTRAMTRILLBRINTENLERGYRGKKKKFRENLYPL FKYLTAINHTLKENLERKFYKGKKKKCHERLYPL FKYLTAINHTLKENLERKFYKGKKKKCHERLYPL FKYLTAINHTLKENLERKFYKGKKKKCHERLYPL			1				GAPAGARGGPAKANSNPFEVKVNROKFOU G
RDKSNYFRDERSYNSNMSPEEKMMKRPA LEQQRHIERKSYNTMEDEETHYGQSLADII KINDIVDSDSDAEDRGTLSGELTAAHFGGG GILHKKTQQGGEREKKPKSRKELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIGTLI SHKTPKSENRDKKEKPKPDAYDMMYRELGF EMKAQPSNRMKTEAELAKEQGEIRLKLEAG RLRRMLGKDEDENVKPKHMSADDIJDGF LDKDDRRLLSYKDGKMNVELOF EMKAQPSNRMKTEAELAKEQGEIRLKLEAG RLRRMLGKDEDENVKPKHMSADDIJDGF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF LDKDDRRLLSYKDGKMNVELOF RSASDAIKFVLRDAMHEMEEMIETKGRALP GLDVLIYLKITGLLFYTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEVVALS QREFELINFLLGILVSAREDVATWQOSSLSLRWA SRLRAPTSTEADHRLSCLAGALLKRCVLM YGSLPSHAMGPILRALLTDHLADCSHPQELO ELCQSTLTEMESQKQLCRPLTCEKSKPPVKL FTRILKVLEFGRKQGSSKEGERKRLHKHK REFKGAVREIKDNOFLARMQLSEIMERDAE RKRKVKQLFFSKLATGEGEWKLALKKRKFKFK RKRKVKQLFFSKLATGEGEWKLKRKFKFKK ASSAGITOVSHCSWPVTYLTLHAVRNVL FKRTFPLKSSSFLSYDKEIPPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHITHCTANWVHLPGTPFDWKIDGPAAAL EVLSSFFFFFLKSSKFNDNIV FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGERTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAFVNL FKRTFPLKSSKFLSTURGEFTRUNLAF FFFYSPLFGFFFTRUSHERGERTLAFVL FRICHTRUNLAFTRUN		ĺ	ĺ			İ	RKTRHDVGLPGVSRARALRKRTOTLLKEVKE
LEQORHHEKKSIYNINEDELTHYGQSLADII KINDIVIDSDSDAEDROTLSGELTAAHFGGG GILHKKTQOEGEREKPKSRKELEELIAKSK QEKRERQAQREDALELTEKLDQUWEIQITLI SIKTRYSSIRNIKKERPKSRYDAYDMMYRELGF EMKAQPSINRMKTEAELAKEEGERILRKLEAE RIRRMILGKDEDENVKRYBHMSADDI.NDGF LDKDDRRLLSYKDGKINIVEEDVOEGOSKEA SDPESINEEGGDSSGGEDTEESDSPDSHI.DLES NVESEEENEKPAKEQROTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSILLGRSMEEG LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVYHLYLCQMFF ESASDAIKFVLRDAMFEMEINETKGRAALP GLDVLIYLKITGLIFTSDFWHPVYTALVCL SQLITKCPILSLQDVVKGLFVCCUFLEFVYALS QRIPPELINFLLGILLYAATPIKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQOSSLSLRWA SRLRAPTSTEANHRILSCLAVGLALLKRCVLM YGSLPSHAIMGPLRALLTDHLADCSHQPELQ ELCQSTLTEMESGKQLCRPLTCEKSKPYPLKL FTPRLVKVLEFGRKQGSSKEGERKRLIHKRK REFKGAVSERRKDOPJCRATLTCEKSKPYPLKL FTPRLVKVLEFGRKQGSSKEGERKRLIHKRK REFKGAVSERRKDOPJCARAMOLSEIMERDAE RRKVVQLFNSLATQEGEWKALKRKFKK KREFKGAVSERRKDOPJCARAMOLSEIMERDAE RRKVVQLFNSLATQEGEWKALKRKFKK FKFFFLVESSSFLSYDKEIPFILIVLKFYLVTLT FSYVK 690 2040 A 5261 1 304 FFFFVLYETGFHHVGQAGLELLTSGPPTW ASQSAGITGVSHCSWPVYVLSTILHAVRNVL FKRTFPLKSSSSLSYDKEIPFILIVLKFYLVTLT FSYVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPFGWKALKARAM FKRTPLKSSSSLSYDKEIPFILIVLKFYLVTLT FYRK FKFFFFFFLKFSYKPONIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV ERVLITFLPAKALLRVACVCRLWRECVRRVLR FRISTFTWSAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADSETFISLEEGCHKRAR KRTSMETALALEKLFFKQCQVLGIVTPGGIVTOF PMGSGSNRPQEIGIGSGFALLFPDIGGIKTOPF HFIKDPKNLTERHOLTEGAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVCADAFR KFFSVPLIEGFFGROGGIGGSFALLEPDGIGKTOPF HFIKDPKNLTERHOLTEGAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVCADAFR KFFSVPLIEGFFGROGGIGGSFGALLEPDGIGKTOPF HFIKDPKNLTLERHOLTEGAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVCADAFR KFFSVPLIEGFFGROGGIGGSFGGLUFPOTGIVTTIFLRKCV EVKDDDLFHSYTTIMALHILGSSK 693 2043 A 5301 362 507 EERKEFEGFGLVTVWVGFTGELDCNRERGILLK ACFPTNIVTLCHSIA KTRAMTRALKNIMHEELLKLKYCHKPKERKFYPL		Į		1		ļ	RDKSNVFRDKREGEYNSNMSPEEKMMKREA
KHNDIVDSDSDAEDRGTILSGELTAAHFIGGGE GILHKKT/QOEGEEREKPERKELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIQTIL SHKTPKSENRDKKEKPKPSRYDMWKEIGF EMKAQPSNRMKTEALELAKEEQEHLRKLEAE RLRRMLGKOEDENVKKPKHIMSADDINDGP LDKDDRRLISYRDGKNUKEDVGEGQSKEA SDPESNEEGGSSGGEDTEESDSPDSHIDLES NWESEEBENERAKEOPGYGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMECG LLVVERIQKCNHPSLAEGNKAKLEKLFGFILL YVGDLATDPPDLTVIDYOKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMECG LLVVERIQKCNHPSLAEGNKAKLEKLFGFILL YVGDLATDPPDLTVIDVYHLYHLQMFF ESASDAIKFVLRDAMHEMEENIETKGRAALP GLOVLIYLKITGLLPF3DEWHPVVTPALVCL SQLITKCPILSIQDVVKGLFVCCLFLEYVALS QRFIPELINFLIGILYIATPNKASQGSTLVHIPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHRIS.SCLAVGLALKRCVLM YGSLPSTHAIMGPLRALITDHLADCSHPOELQ ELCQSTLTEMESQKQLCRPLTCESKPVPIKL FIPPLIV KV LEFGRKQGSGELDISGGDPTW ASQSAGITGVSHCSSKEGEKSKPVPIKL FIPPLIV KV LEFGRKQGGELLTISGDPFTW ASQSAGITGVSHCSSKEGEKSKPVPIKL FRYTPILSSSFLSYDKEIFPILIVLKFYLVTLT FSVK FSYTPLIVELGELTISGDPFTW ASQSAGITGVSHCSSFPSHVIVLKFYLVTLT FSYVK SPILLSSFFFFLKSYKPQNIV G91			J	J			LEOORHHEKKSIYNI NEDEFI THYGOSI ADIF
GILHKKTQQEGEREKPKSRKELEELIAKSK QEKRERQAQREDALEL TEKLDQDWKEIQTLI SHKTPKSENRIDKKEKPKPSDAYDMMVRELGF EMKAQPSNRMKTEAELAKEEQBHLRKLEABE RLRRMI.GKDEDENVKKEKPKPDAYDMMVRELGF EMKAQPSNRMKTEAELAKEEQBHLRKLEABE RLRRMI.GKDEDENVKFKHMSADDI.NDGF LDKDDRRLLSYKDGKNNVEEDVÇEQSKEA SDPESNEEGDSSGGEDTESDSPDSHI.DLES NVESEEBERPAKEQROTPGKGLISGKERAG KATRDELPYTPAAPES STEERSILLGRSMEEG LLVVERIQKCHPSLABGNKAKLEKLFGFILE YVGDLATDDPDLTVIDKLVYHLYLCQMET ESASDAIKFVLRDAMEMEMENIETKGRAALP GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEFVVALS QRIPBELINFLLGILVYARGDAYDQSSLSLRWA SRLRAPTSTBANHRLSCLAVGLALLKRCVLM YGSLSFRHAMGPLRALLTDHLADCSHQELG ELCQSTLTEMESGKQLGRHTCEKSKPVPLKL FTFRLVKVLEFGRKQGSSKEGERRLIHKHK REFKGAVSERIKNDNFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGBWKALKRKFKK FKK FFFFVFLYETGFHHVGQGLELLTSGDPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKTTPLKSSSFLSYDKEIFPILVLKFYLVTLT SFVK 690 2040 A 5261 1 304 PFFFVFLVETGFHHVGQGLELLTSGDPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKTTPLKSSSFLSYDKEIFPILVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTTCTANWVHLPGTFPGWKIDGPAAAL EVLSSFFFFFLKFSYKPONIV FKTTPLKSSSFLSYDKEIFPILVLKFYLVTLT SFVK 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLSTFFFFFLKFSYKPONIV FRINDFRINLTGHAMADSETTISLEEGCHKARA KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGSGFALLFPQIGGIKIQPF HFIKDPKNLTLERHQLTECKHKARKFKK FKKFTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGSGFALLFPQIGGIKIQPF HFIKDPKNLTLERHQLTECHJDAPPELRVVLV FGYNCCKVGASNYLQVVSTTSDMNIILAGG QVDNLSSLTSEKNPLDIDAGSVCGSFSGHRA KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGSGFALLFPQIGGIKIQPF HFIKDPKNLTLERHQLTECHJDAPPELRVVLV FGYNCCKVGASNYLQVVSTTSDMNIILAGG QVDNLSSLTSEKNPLDIDAGSVCGSFSGHRA FKFPSVPLIEFFFFFFNGFFGFGGRGGGTALLFPQIGGIKIQPF HFIKDPKNLTLERHQLTEGSFSGHRA FKFPSVPLIEFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF		1				}	KHNDIVDSDSDAEDRGTI SGELTA AHEGGGG
QEKRERQAQREDALELTEKLDOWKEIGTIL SHKTPKSENRIOKKERPKPDA YDMMYRELGF EMKAQPSNRMKTEAELAKEEQEHLRKLEAE RLRRMLGKDEDENVKRYRHMSADDLNDGFT LDKDDRRLSYKDGKNNVEEDVQEEQSKEA SPPESNEEGDSSGGEDTEESSSPSHILDLES NVESSEEBKPAKERQFOTPOKGLISGKERAG KATRDELPYTRAAPESYEELRSLLLGRSMEEQ LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVYHLYHLCQMFF ESASDAIKFVLRDAMHEMEENIETKGRAALP GLDVLYIKLKITGLLPTSDFWHPVVTTALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEVVALS QRFIPELINFLIGILVATPSWAYDEVTTALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEVVALS QRFIPELINFLIGILVATPSWAYDEVTTALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEVVALS QRFIPELINFLIGILVATPSWAYDEVTTALVCL SQLLTKCPILSLQDVVKGLFVCCKSKYPVLKI VSGLPSTHAIMGPLRALITDHLADCSHPQELQ ELCQSTLTEMESQKQLCFPLTCESKPVPLKI FIFRLVKVLFGRKQGSSKEEQERKRLIHKEK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDMQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK REFKTPELLSSSSLSYDKEIFPILVLKFYLVTLT SFVK S991 2041 A 5270 3 158 NCHTTHCTANWVILLPGTPFGWKIDGPAAAL EVLSSFFFFPLKFSYKRQNIV GWENNESSLSTEVRFJDIOAGNAAA FFFFVELTFAALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFFKQCQVLGIVTIGIVTIGVTV PMGSGSNRPQEIEIGESGFALLFPQIGIKIQPF HFIKDFKNLTERHQLTEVGLIDDPELRVVLV FGYNCCKVGASNYLQVVSTTSSDMIILAGG QVDNLSSLTSENPHJDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVEADNFR KFFSSYLIFGFFGNGEIGEGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGGRALLFROCKPADHR SECKEFFGFGFGNGEIGEGGGRALLFROCKPADHR KFFSSYLIFGFFGNGEIGEGGGRALLFROCKPADHR KFF				1			GLI.HKKTOOFGFFREKPKSRKELIEFI LAKSK
SHKTPKSENRDKKEKPRPAYDMMYRELGE EMKAQPSNRMKTSALAKEGGEHLEKLEAE RLRRMLGKDEDENVKKPKHMSADDLNDGFV LDKDDRRLLSYKDGKMNVEEDQEEGDSKED SPESNEEGGDSSEGGDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPEGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMEEG LLVVERIQKCNHPSLAEGNKAKLEKLFGFILE YVGDLATDDPPDLTVIDKLVYHLYHLCQMFF ESASDAIRFVLRDAMHEMEMIETKGRAALP GLDVLIYLKITGLLFPTSDFWHPVTPALVCLL SQLLTKCPHI SLQDVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLRPTLCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNGFLARMQLSEIMERDAE RKRVKQLFNSLATDGEGWKALKRKKFK FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNGFLARMQLSEIMERDAE RKRVKQLFNSLATDGEGWKALKRKKFK FKRTPPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5261 1 304 FFFFVLYETGFHHVGQAGLELLTSGDPPTW ASGSAGTGVSHCSWIVYLVSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADAVCVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADAVCVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADASVQLUSTFTSISLECRGRIKRAR KRTSMETALALEKLFPKQCQVLGIVTFGGVYT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQFF HIKDPRNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVSTFSDMIILAGG QVDNLSSLTSEKNPLDIDASGVVGLJSSFGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVEADAFR KFFSVPLFGFFGNGGICGDRIVTGNFLRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLUTYVYGFTQELDCNRERGILLK CFPTNIVTLCHSIA 694 2044 A 5310 1 204 RYLTAINHTLCHLKKFKKKKKKPKLLDRFK KTRANKRLNMHEENLKTKKQHRKERLYPJL	`}	1	ļ			1	QEKREROAOREDALELTEKI DODWKEIOTTI
EMKAQPSNRMTTAABLAKEGBHLRKLEAE RIRRMLGKDEDENVKPKHMSADDLNDGFY LDKDDRRLSYKDGKMNVEEDVQEQSKEA SDPESNEEGGDSGGGDTEESDSPDSHLDLES NVESEEBNEKPAKEQRQTPGKGLISGKERAG KATTDELPYTFAAPESYEELRSLLGRSMEEG LLVVERIQKCNHPSLAGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVYHLYHLCQMFF ESASDAIKFYLRDAMHEMEEMIETKGRAALP GLDVIJLKITGLLFPDFWPPVYTPALVCL SQLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFPELINFLLGILYIATTPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRAH YGSLPSFHAIMGPLRAH YGSLPSFHAIMGPLRAH REFKGAVREIRKNDGLARMQLSEINBRDAE RKRKVKQLFNSLATQEGEWKALLRKKFKK REFKGAVREIRKNDGLARMQLSEINBRDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK FFFYFLVETGFHHVGQAGLELLTSGDPFTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK 691 2041 A 5270 3 158 NCHTTFLCTANWVHLPCTPPGWKIDGPAAAL EVLSSFFFFELKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECGESSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVL YMADSETFISLEECRGHKRAR KRISMETALALEKLFPKQCQVLGIVTFGIVVT PMGSGSNRPQEEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERIQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFYTYARGGNVFLACH GSATVLLNEDVSDEKTAEAAMGRLKAANIPE INTIGFMFACVGRGFTENLOTKKKTKANTERCHLYKFKYGKKDKKFKFLLFRF GYTATAMTLKKRINHTERENLKTKKYGHRKERLYPL				ļ			SHKTPKSENRDKKEKPKPDAVDMMVREIGE
RILRRMLGKDEDENVKKPKHRSADDIJADDEN LDRDDRRILSYKDGKMNVEEDVQEGSKEA SDPESNEEGDSSGGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTTAAPESYEELRSILLIGRSMEEG LLI-VVERIQKCNIPSI.AEGNKAKLEKLFGFLLE YVGDLATDPPPDLTVBL-VHLYHLCOMFF ESASDAIKFVLRDAMHEMEEMIETIGGRAALP GLDVLIYLKITGLLFPTSDFWHPVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTL-VIPPER ALGKNESELL-VVSARGVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESSQC.RPTLCEKSKP-VPLKL FTPRL-VKVLEFGRKQGSSKEE-GERKRLIHKRK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKKVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRKVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRKVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRKVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRKVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRYVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRYVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE RKRYVKQLINSLATGBGEWKALKKKKKK REFKGA-VREIRKDNQFLARMQLSEIMERDAE REFKTPFLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK 691 2041 A 5270 3 158 NCHTTCANWVILPGTPPGWKIDGPAAAL EVLSSFFFFFLKSSKFDNNILAGV ERVLTFLTAMAUHLLPGTPPGWKIDGPAAAL EVLSSFFFFFLKSSKFDNNILAGV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADGEGGECRISTSISLECRGHKRAR KRISMETALALEKLFPKQOVLGIVTFGGIVT PMGSGSNRPQEEIGESGFALLFPQIEGIKIOPF HFIKDPKNLTLERHQLTEVGLUDNPELRVVLV FGYNCCKVGASNYLQOVSTFSDMNILAGG QVDNLSSTSEKNPLDIDASGVVGLSFGHRI QSATVILNEDVSDEKTARAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGGICGRUTTGNFLRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFRGFGLUTYVYGFTQELDCNRERGILLK ACFPTNIVTLCHIA 694 2044 A 5310 1 204 RVLTANHTIKENLRKFYKGKKKKKPKPLLRFPL	i	ĺ	l	1 1			EMKAOPSNRMKTEAELAKEEOEHI RKI FAF
LDKDDRRLISYKDGKMNVEEDVQEEQSKEA SDPESNEEGGSSGEDTEEDSDPSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTHAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPAAPESYEERSLLIGRSMEEG LLVVERIQKCHIPATOPPOLIVIDKI, VVHLYHLCQMFF ESASDAIKFVLRDAMHEMEEMETKGRAALP GLDVLIYLKITGLIFTSDFWHPVVTPALVCLI SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLIGLIAPTHSDFWHPVVTPALVCLI SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLIGLIAPTHSDFWHPVVTPALVCLI SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLIGLIFTSDFWHPVVTPALVCLI SQLLTKCPILSLQDVVKGLFVCLFLEXSKFVPILSL GRFBELINFLSCLAVGLALLKRCVLM YGSLSPSHAIMGPLAATHBOSSGTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAFISTEANHIRLSCLAVGLALLKRCVLM FFFFYLVETGFHALLTDHLADCSHOPELO ELCQSTLTEMESQKQLCRPLTCEKSKFVPILKLI FTPRLVKVLEFGRKQQSSCHLERKEDKSHRENDA RKRKVKQLFNSLATQEGEWKALKRKKFKK REFKGAVREIRKDQGTLELTISGDPFTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FRKTFPLLSSSSLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCANWVILLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 NCHTTHCANWVILLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 693 COMPANIENT OF THE STANDARD	ł	l	l	i i			RLRRMLGKDEDENVKKPKHMSADDLNDGEV
SDPESNEEGGSSGGEDTESDSPDSHLDLES NVESSEENERPAKFOROTPGKGLISGKERAG KATRDELPYTFAAPESYEERSLLICGRSMEEC LLVVERIQKCNHPSLAEGGRAKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVYHLYHLCOMFF ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLLFYTSDFWHPVYTPALVCL SQLLTKCPHLSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGIL YIATPIKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCOSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQEKRLIHHKR REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKKKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKKKFKK RKKRKKFKK RSKSGAVREIRKDNQFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKKFKKK RKKRFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKKFKKK RFFFFLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKSSYKPONIV 692 2042 A 5282 56 1268 GMEPVGCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR HTHSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLIPHTVLYMADSETFISLEECRGHKRAR KRTSMETAALLEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPGIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQVVSTFSDHMIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVFADAFR KFFPSVPLFGFFGNOEIGCDRIVTGNFLRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKKKKPKLPLPLPK KTRAMRRILNMHEENLKTKKQHRKERLYPPL KTRAMRRILNMHEENLKTKKQHRKERLYPPL KTRAMRRILNMHEENLKTKKQHRKERLYPPL	Į						LDKDDRRLLSYKDGKMNVEEDVOFFOSKFA
NVESEEENERPAKEOROTPGKGLISGKERAG KATRDELPYTRAPESYEELRSLLLGRSMEEC LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVYHLYHLCQMFF ESASDAIKFYLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYYVALS QRFPELINFLLGILYHATPNKASQGSTLVHPFF ALGKNSELLVVSAREDVATWQQSSLSLRWA SRIRAPTSTEANHISCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRGSSKEFQERKRLIHHKR REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK FFPTVFLVETIGFHVQQAGLELLTSGDPFIW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIPFILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANVVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKSFSFKPONIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTTLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADSETFISLECCRGHKRAR KRTSMETALALEKLFFKQCVJLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPGIGIKIQPF HFIKDPKNLTLERHQLTECQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPGIGIKIQPF HFIKDPKNLTLERHQLTECQVLGIVTPGIVT PMGSGSNRPQEIEIGESGFALLFPGIGIKIQPF HFIKDPKNLTLERHQLTECQVLGIVTPGIVT PMGSGSNRPQEIEIGESGFALLFPGIGIKIQPF HFIKDPKNLTLERHQLTECQVYGLDNPELRVVLV FGYNCCK VGASNYL QQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGVYYAKGNVEADAFR KFFPSVPLFGFFGNGEIGCCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLUYWYGFGQELCDNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKKKPLPLDLRPK KTRAMRRILMHEEMLKTKKQHRKERLYPL							SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
KATRDELPYTFAAPESYELRSILLIGRSMEEG LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE VYGDLATDPPDLTVIDKLVVHLYHLCQMFF ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKIGLLFTSDFWHPVTPPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPMKASQGSTLVHPFR ALGKNSELLVAREDVATWQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSEGERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK ASAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILVLKFYLVTLT SFVK	1		ĺ				NVESEEENEKPAKEOROTPGKGLISGKERAG
LLVVERIQKCNIPSLAEGNKAKLEKLFGFLLE	ŀ						KATRDELPYTFAAPESYEELRSLLLGRSMEEO
YYGDLATDDPPDLYUDKLVYHLYHLQQMFF ESASDAIKFYLRDAMHEMEEMIETKGRAALP GLDVLIYLKTGLLFFTSDFWHPVYTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYYALS QRFIPELINFLGII YAITPIKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRISCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEQERKRLIHKHKK REFFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK FFFFVLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIVYLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK SPVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV FGVTATALALEKLFFKQCVLGIVTPGVVTAEEL ENVRILPHTVLYMAADSTFISLEECRGHKRAR KRTSMETALALEKLFFKQCQVLGIVTPGIVT PMGSGSNRPQEIEJGESGFALLFPQIEGIKJQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVUL FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INNTGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGGLVIYWYGFJQELDCNRERGILLK ACPPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRFK KTRAMTRRLNMHEENLKTKKQHRKERLYPL							LLVVERIOKCNHPSLAEGNKAKLEKLEGFLLE
ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLIFPTSDFWHPVVTPALVCL SQLITKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFILGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRVKQLFNSLATQEGEWKALKRKKFKK FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRVKQLFNSLATQEGEWKALKRKKFKK RFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITOVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK SPV							YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP
GLDVLIYLKITGLIFTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLIGILYIATPNKASQGSTLVHPPR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLH YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFFGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK REFFGAVREIRKDNGFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK REFFFVLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTPPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFLKFSYKPONIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKUPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMVIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRFK	ľ			{ i			ESASDAIKFVLRDAMHEMEEMIETKGRAALP
SQLLTKCPILSLQDVVKGLFVCCLFLEYYALS QRFIPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLSPSHAMMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKLL FTPRLVKVLEFGRKQGSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEGWKALKKFKKK ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHITTCTANWVHLPGTPPGWKIDGPAAAL FYRSTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADSETFISLECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQEIGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNVLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE INTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFFLRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLKKFYKGKKDKPLDLRPK KTRAMRRINMHEENLKTKKQHRKERLYPL	l			1			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
QRFPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RRKKVKQLFNSLATQEGEWKALKRKKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RRKKVKQLFNSLATQEGEWKALKRKKFKK ASQSAGITGVSHCSWPVIYVLSTLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK SPVK SPVK SFVF	İ						SOLLTKCPILSLODVVKGLFVCCLFLEYVALS
ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQEKRKLHKKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE FREKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCGGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCK VGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK 694 2044 A 5310 1 204 RVLTANHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL							QRFIPELINFLLGILYIATPNKASOGSTLVHPFR
SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESOKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	1			1 1	1		ALGKNSELLVVSAREDVATWOOSSLSLRWA
YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL	}						SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVL YMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 I 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL				1 1			YGSLPSFHAIMGPLRALLTDHLADCSHPOELO
FTPRI_VKVLEFGRKQGSSKEFQERKRLIHKHK				1			ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL
REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVQQAGLELTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQUEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNVLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL		, ,		1 }			FTPRLVKVLEFGRKQGSSKEEOERKRLIHKHK
RKRKVKQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL] }			REFKGAVREIRKDNQFLARMQLSEIMERDAE
690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SPVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAELE ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCK VGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVTWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL							RKRKVKQLFNSLATQEGEWKALKRKKFKK
ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	690	2040	Α	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIHLAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL		' l		} }			ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
SFVK	j	J		}	j		FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL							
692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK FYSVPLFGFFGNEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	691	2041	A	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVTV PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK FYSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL					İ		EVLSSFFFFFLKFSYKPONIV
ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVL YMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCK VGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	692	2042	A _	5282	56	1268	
THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVL YMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCK VGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL					l		
ENVRILPHTVL YMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCK VGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	l		ŀ		ı		THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	j]	ļ		}	j	ENVRILPHTVLYMADSETFISLEECRGHKRAR
PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	1						KRTSMETALALEKLFPKOCOVLGIVTPGIVVT
HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	1	1	[1	[ſ	PMGSGSNRPQEIEIGESGFALLFPQIEGIKIOPF
FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 I 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	ł	1	ļ	- 1	1		HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	l	Į	İ	-	ļ		FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	j		ļ	1			QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	j	l	l	1			QSATVLLNEDVSDEKTAEAAMORLKAANIPE
KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRLNMHEENLKTKKQHRKERLYPL	ł	1	ì	-	ł		IINTIGFMFACVGRGFQYYRAKGNVEADAFR
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRLNMHEENLKTKKQHRKERLYPL	I	j	J	J			KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRLNMHEENLKTKKQHRKERLYPL							EVKDDDLFHSYTTIMALIHLGSSK
694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	693	2043	A]	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
KTRAMRRLNMHEENLKTKKQHRKERLYPL							ACFPTNIVTLCHSIA
KTRAMRRRLNMHEENLKTKKQHRKERLYPL	694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
RKVAAKA	i	j	- 1]	ļ	KTRAMRRRLNMHEENLKTKKQHRKERLYPL
			1				RKYAAKA
695 2045 A 5315 125 1596 ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP	695	2045	A	5315	125	1596	ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

nucleotide corresponding to corresponding to corresponding to corresponding to the correspond	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence 99/496 corresponding 1 state and in section 1 state			1100				
Sequence 99/496 correspond neid residue of peptide neid residue of peptide sequence sequence sequ			İ		t .		
amino acid residue of peptide sequence per per per per per per per per per pe	1			1	correspondi		
	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Peptide			ļ	1	amino acid	of peptide	
			ł			sequence	, ,,,,
TOSCKEEGPAPCKOMKLEAAGGPSALNFDSP SISLESLISHKITETREKPEKOPKULLORDDPA							
SSLFESLISPIKTETFKEPWEGKPLLIQREDPA					sequence		
LATYYGSLFKLTDLKSLCSRGMYYGRDVDK			İ				
CRCVNICKKVILNKDGKAHFLQLRKDFDQKY ATIQPHQPQRFREDEL WRIGEKLECYTGSLYGS NVYTTPAGSGQLPPHYDDVEVFILQLEGEK WRLYHFTVPLAREYSVAEREIGERYHEFEM KPGDLLYPFGTHQADTPAGLAHSTHVTIST VQNINSWGDFLLDTSGLVPTTAGEDVELRTG IPRQLLLQVESTTVATRRLSGFLRTLADRLEG TKELLSSDMKKDFIMHRLIPYSAGDGALST: OKKRLDSVRLQKFGBINLYLLPDQDQSD ETQEKMYVIYHSLKNSRETHMMGNEETEFF GLRPFSHLDALKQWNSPAISVKLKLTTED GREENLYLSLWTECLIQVV EKESLVLSLWTECLIQVV OKKRLDSVRLAKTINGNESTHAMMGNEETEFF SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN VEKLQVLLNCMTERTYQPKKDKAERKLAYN EEQHIKFDKQKLYYHATKAMTHFTDECVKK YGAFLNSSEEWIRKMHLRKQLSLSTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPHYPSSNTLVEMTLGMKLLSSENGCHT MTPHYPSSNTLVEMTLGMKLLSSENGCHT MTPHYPSSNTLVEMTLGMKLLSSENGCHT MTPHYPSSNTLVEMTLGMKLLSSENGCHT MTPHYPSSNTLVEMTLGMKLLSSENGCHT MTPHYPSSNTLVEMTLGMKLLSSTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPHYPSSNTLVEMTLGMKLLSSTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPHYPSSNTLVEMTLGMKLLSSTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPHYPSSNTLVEMTLGMKLLSSTNQCFDI EEVSKYQETYDELQETLPQKMPTASSGICHT MTPHYPSSNTLVEMTLGMKLLSSTNQCFDI EEVSKYQETYDELQETLPQKMPTASSGICHT MTPHYSSNTLVEMTLGMKLLSTNQCFDI EEVSKYQETYDELQETLPQKMPTASSGICHT MTPHYSSNTLVEMTLGMKLLSTNQCFDI EEVSKYQETYDELQETLPQKMPTLGGTAGGSG LAKTYTGSLERNIVGLG A 5324 266 714				1			
ATIQPHQPQRFKDELWRIQERICTYGSLYGE WRI_YHPTYPAGSGQ_PPHYDDVEVFILQLEGEKH WRI_YHPTYPLAREYSVEAERIGRPVHEFM: KPQDLI_YPPRGTHIOAOTIPGGLAHSTHYTIST YQNINSWGDFLLDTISGL_VPDTAKEDVELRTG IPRQLLLQVESTTVATRALSGFLRTLADRLEG TRELLSQMKKOFTMHRLPYSAGDGABLSTP GGKLPRLDSVVRLQRKDHIVLTVLPDQDQSD ETQEKMYYTYHSLKNSRETHMRGREETEFH GLRPPLSHLDALKQIWNSPAISVKDLKLTTDE EESSL_VLS.WTECLQVV LMKXYLEAAELGEISDHIKLLRLSSSQGTLTDE EESSL_VLS.WTECLQVV LMKXYLEAAELGEISDHIKLLRLSSSQGTLTDE EEQHKPDKQKL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EEQHKPDKQKL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQKL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQKL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQKL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNKSEEWIRKMH.HLRQLLSLTNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNCOPT EECHKPDKQCL_VYHATKAMTHFDECVKK YEAFLNCOPT EECKWYQCLSTDEGTHQCRQVKCLLATTYPHPSTSTLAGGLALDTATASSCC EECKWYQCLSTDEGTHQCRQVKCLLATTYPHPSTSTLAGGLALDTATASSCC EECKWYQCLSTDEGTHQCRQVKCLLATTYPHPSTSTLAGGLALDTATASSCC EACHT CHARLES EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYYTHILLAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYQCLATTAGA EECHLAND EECKWYGCLATTAGA EECHLAND EECKWYGCLATTAGA EECHLAND EECKWYGCLATTAGA EECKWYGCLATTAGA EECHLAND EECKWYTHTAGA EECKWYTHTAGA EECKWYTHTAG							
NVYITPAGSGGLPPHYDDVEVFILQLEGEK WRLYHTYPLARESVSEAERIGRYHEFM. KPGDLLYPRGTIHQADTAGLAHSTHVITS VPANSWGDFLLDTSGLYPTDAGLAHSTHVITS VPANSWGDFLLDTSGLYPTDAGLAHSTHVITS TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG TREILSSDMKKDFIMHRLIPYSAGDAGLEG GEVEKNYLIYHSILKNSRETHIMMGREEFTEH GIRPPISHLDALKQWINSPAISVKDLKLTTDE ERESLVIS.WTECLIQVV LMKXYLEAAELGEBOHIMLIRISSSQGTET SLQDIDSRLSPGSSLADAWAHQEGTHPKDRIV SLQDIDSRLSPGSSLADAWAHQEGTHPKDRIV VEKLQVILINGHETYQPKKNDAAERILAN VEKLQVILINGHETYQPKKNDAAERILAN VEKLQVILINGHETYQPKKNDAERILAN VEKLGVLINKSEWIRKHMI HLR KOLISTNOCPD EEVSKYQEYTNELQETLYQKMFTASSGIKHT MTIPYPSSYTLVEMTLOMKSLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNDCI.NTOCPD EEVSKYQEYTNELQETLYQKMFTASSGIKHT MTIPYPSSYTLVEMTLOMKSLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNDCI. 697				Ì			
WRLYHPTVPLAREYSVEAERRIGRPVPLEFML KPGDLLYPREAGLAISTHVTIST YQNNSWGDFLLDTISGLVFDTAKEDVELRTG IPRQLLLQVESTTVATRILSGTALADILEG TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP GGKLPKLDSVVPLQFKDHIVLTVLPQDQSD ETQEKNVYYINISKASRETHMGNEEFEFH GLRFPLSHLDALAQIVNSPAISVKDLKLTTDE EKSELVLSLWFILKSKASRETHMGNEEFEFH GLRFPLSHLDALAQIVNSPAISVKDLKLTTDE EKSELVLSLWFILKJECLIQVV LMKXYLEAAELGEISDHTKLLRLSSSQGTIET SLQDIDSRLSPGGSLADAWAHQEGTHFKDRN VEKLQVLINCMTEIYYQFKKDKAERRLAYN EGHIKFDKQKLYYHATKAMHTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNOCFD EEEVSKYQEYTMELQETLPGKTSSQGKTH MTPIPYPSSNITLVEMTLGMKKLKEEMEGVVKE LAENNHILESGSLTMDGGLGSSLTMDGGLGSTANDGGLGVKE LAENNHILESGSLTMDGGLGSSLTMDGGLGSTANDGCL GPP PPGFKQFSCLSLFSSWDYRHLPPHLANFSREG VSPSWFGWSRTPDFR LPIRKSLRSVRSGFFTSQSFTRNLDGTASSGC LAKTVTGSLFRINVQLGLLVAGGIGALLGTP VGGLLMAFQKYSGETVQERKQKDKALHEL KLEEWKGRLQVTEHLPEKEISSLQEDEPENDA KKIEALINLPRNPSVIDKQDKD KKIEALINLPRNPSVIDKQDKD KKIEALINLPRNPSVIDKQDKD KKIEALINLPRNPSVIDKQDKD KKIEALINLPRNPSVIDKQDKD KKIEALINLPRNPSVIDKQDKD CFGAESSVFETFVQKKQKGIKATTIVCPFFIKTGM FGGCTTGCPSLIPLBFKYVAFCALAQVLG LHAMDGFADOK PTAEEMSSLTFESSPFLAKRSWFGNPTSLDKEE QIFLVLKDKPLSSKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGSVFCKFVRFQVDISSSEG PPSPSRBGSGGGGIYSVTFTLISGFSRRFKRV VETIQAQLISTHDQPSVQALADEKNGAQTTRP AGAPPRSLQPPFGRPDFDESSSPRGFPKDKK LIATNGTEL GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHMIIERIKV LADVKLVAJPGPGRTJLFEFSPGTTTTLE GSSSVPCIFCEEHPPVAGQDKLLKHKSELQ KQRLRILEQQQQFGNTDTSTKTHLDHDMR KKQHRRILEQQQQFRNTDTSTKHLKDHMR KKQHRRILEQDQGCRNTDTSTKHLKDHMR KKQHRRILEQDQGCRNTDTSTKHLKDHMR KKQHRRILEQDQGCRNTDTSTKHLKDHMR KKQHRRILEQDQGCRNTDTSTCHATLADHUR KKQHRRILEQUBCHTFOTT TYENDTILLWTLSDSESSULTAQUELYYFF TYENDTILLWTLSDSESSULTAQUELYYF							
KPGDLLYPPRGTHRQADTFAGLAISTHVTIST				ļ			1 '
YQNNSWGDFLLDTISGLVPDTAKEDVELRTG IPRQLLLQVESTYVATRLS GTALADREG IPRQLLLQVESTYVATRLS GTALADREG TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP GGKLPKLDSVVRLQFKDHIVLTVLPQDQSD ETQEKMVYIYHSLKNSRETHMMGNEETEFH GLRFPLSHLDALAQIWNSPAISVKDLKLTTDE EKSEJ.VLSLWTECLIQVV	1	İ		<u> </u>		Ì	
IPROLLLOVESTIVATRRISGERITLANDEG IKKLISSIMKKOPHMIRLIPYSAGOGABLSTP GGKLPRLDSVVRLOFKOHIVLTVLPDODQSD ETQEKMVYIVISLKYSRETHMMGNEETEFH GLRPRISHLDALKQIWNSPAISVKDLKLTTDE EKESLVLSLWTECLIQVV				•			
TKELLSIDMKKDFIMHRI,PPYSAGDGAELSTP GGKLPRLDSVVRL,GFKDHVLTVLPDQDQSD							! -
GGKLPRLDSVYRLQFKDHIVLTVLPQDQQSD ETQEKMYVIYHSLKNSRAISVKDLKLTTDE ETQEKMYVIYHSLKNSRAISVKDLKLTTDE EKESLVLSLWTECLIQVV							; · · · ·
			ł				i i
GLRPPLSHLDALKQHWNSPAISVKDLKLTTDE		1	l				
	1						
696	l		İ				
SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN	696	2046	A	5318	1476	742	
EEQHKFDKQKLVYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGBGLRNVDCL 697 2047 A 5320 244 478 LDYNFFLEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTIPDFR 698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPTTRLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKLESSLQEDEPENDA KKIEALINLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVPVETTVQKQKGIKTTIVCPFPIKTGM FEGCTTGCFSLLPLIEPKYAVEKIVEALIQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGSGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPFRPPELSSSPRRGPPLDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLLEPLSLPESPGTTTLE GSPSVPCIFCEEHFPVAEQDKLKHMIEHKIV IADVKLVADFQRYILYWRKFFTEGPTDFCSV IRINSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQCERNDTNFHOVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CITLQKKLDNLQCLYCEKTFROKNTLKDHMR KKQHRKINPKNREVRPTYNDVLLICHSWEEL LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CITLQKKLDNLQCLYCEKTFROKNTLKDHMR KKQHRKINPKNREVRPTYNDVLLIGLGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIREDRYVINNTLEJGKSWEE LLFRYGNYELLDRYCHOPCCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYPNDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKSSESULTALQEQNENVPIISE							
SEAFLNKSSEWIRKMLHLRKQLLSLTNQCPDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVYKE LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMDGGLRNYDCL LAENNHILESGGSLTMAGGRYSGGVYHLPPHLANFSREG VSPSWPGWSRTPDFR VSPSWPGWSRTPDFR VSPSWPGWSRTPDFR LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERCKQLRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQD							
BEEVSKY QEYTNEL,QETLPQKMFTASSĞİKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNVDCL	1						EEQIHKFDKQKLYYHATKAMTHFTDECVKK
MTPIYPSSNTLVEMTLGMKKLKEEMEGGVVKE LAENNHILESGGSLTMDGGLRNVDCL		ľ					YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
LAENNHILESGGSLTMDGGLRNVDCL							
697 2047 A 5320 244 478							1 .
PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR 698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETTVGERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGYNGLADYCASKFAA FGFAESVFVETVQKQKGKIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKVPRFQVDISSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPSLQPPGRPDPELSSSPRRGPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGTYGDSKQALADEKNGAQTRP AGAPPSLQPPGRPDPELSSSPRGPPTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTILQKKLDNLQCLYCEKTFROKNTLKDHMR KKQHKINPKNREYDRRYVINYLELGKSWEE VQLEDDRELDDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
VSPSWPGWSRTPDFR	697	2047	A	5320	244	478	1
698 2048 A 5324 266 714 LPIRK\$LR\$VR\$GFFT\$Q\$PITRNLDGTA\$G\$C LAKTVTG\$LPRINV\$GLR\$GLVAGGIIGALL\$GTP VGGLKMAPQKY\$GETVQERKQKDR\$KALHEL KLEEWKGRLQVTEHLPEKIES\$LQEDEPENDA KKIEALINLPRNP\$VIDKQNKD 699 2049 A 5334 699 277 RPHGHLVCIS\$SAGL\$GVNGLADYCA\$KFAA FGFAE\$VFVETFVQKQKGIKTTIVCPFFIKT\$GM FEGCTTGCP\$LLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLK\$FLPLKTGLILADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEM\$SLTPE\$SPELAKR\$WFGNFI\$LDKEE QIFLVLKDKPL\$SIKADIVHAFL\$IP\$L\$SH\$VL\$S QT\$FRAEYKA\$SGGF\$VFQKPVRFQVDI\$S\$SEG PE\$P\$PRDG\$GGGGIY\$VTFTLI\$GF\$RR\$FKV VETIQAQLL\$THDQP\$VQALADEKNGAQTTP AGAPPR\$LQ\$PF\$GRDP\$DEJ\$S\$PRR\$GPFKDK\$KLLATNGTPL 701 2051 A 5346 3 1383 HA\$VLFCRVMAA\$KTQGAVARMQEDRDG\$C \$TVGGVYGD\$KDCLL\$PL\$L\$PE\$FGGTTILE G\$P\$VCIFCEEH\$PVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWKR\$FTEQPITDFC\$V IR\$N\$TAP\$FEQENYFLLCDVLPEDRILR\$ELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEP LGN\$SVILNHMAREHAFNIGLDVNICNEEL CTTLQ\$KLDNLQ\$LYCEKT\$FDKNTLKDHM\$KKQHRKINPKNREYDR\$FVVINYLELGK\$WEE VQLEDDRELLDHQEDDW\$DWEEH\$PA\$AVCL FCEKQAFTIELLVHMEDAHEPDLLKIKSELG LN\$FYQQVKLVNFIRRQVHQCRCYGCHVKFK\$ KADLRTHMEETKHT\$SLPDRKTWDQLEYY\$P\$ TYENDTLLWTLSD\$SEDLTAQEQNENVPIISE DT\$KLYALKQ\$\$SULNQLLL\$	1						
LAKTYTGSLFRINVGLRGLVAGGIIGALLGTT VGGILMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTIFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFVAEQDKLLKHMIIEHKIV IADVLLVAMFORYILLYMRKFTEQPITDFCSV RINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQCERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	(00	2010					
VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD	698	2048	Α	5324	266	714	·
RLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRRPSVIDKQDKD						ļ	
KKIEALLNLPRNPSVIDKQDKD							
699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTIFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHFASAVCL PCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFTRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSILPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFTRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	699	2049	A	5334	699	277	
FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEFSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL 701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEP LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINVLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		20.5	**	3331	0,7	2,,	
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL 701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEP LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							, , ,
LHAMDGFADQKK							· • •
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	700	2050	Α	5344	3	614	
PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
AGAPPRŠLQPPPGRPDPELSSSPRRGPPKĎKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
TOTAL TO THE PROPERTY OF THE P	1						
701 2051 A 5346 3 I383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLEIGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
STVGGVGYGDSKDCILEPLSLPESPGGTTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		2077					
GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	701	2051	Α	5346	3	1383	
IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1					;	
LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL] .				·		
CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL]						, , , , , , , , , , , , , , , , , , , ,
KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
DTSKLYALKQSSILNQLLL	[
	702	2052	Α	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRHIIVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK
703	2053	A	5380	278	657	TVASRIAFFL LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MVTLFQMWVVPLYFTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCPYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI S
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	Ā	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVOEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFIIPPV SAEIIRKMQQ
711	2061	Α	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749 .	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, =possible nucleotide deletion, \=possible nucleotide insertion KAPELLQGQSEDEQPDASQMHVYSLGMTLY
						WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
713	2063	Α	5506	22	478	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPLHYMVIMN PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR CGHHEVDHFLCEMPALIRMACISTV
714	2064	Α	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR RLTSSVSCALDEAAAALTRMRAESTANAGQS DK
715	2065	Α	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM GRTALFHHSGGSSGYESLRRDSEATGSASSAP DSMSESGAASPGARTRSLKSPKKRATGLQRR RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG SLKEPFEIK YEIDDVERLQRPRPTPREAPTQG LACVSTRLRLAERRQQRLREVQAKHKHLCEE LAETQGRLMLEPGRWLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVAASAAIPGPQEVDV
716	2066	Α	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD SSSSERTMDLVLEMCNTNSIHWCGISGRQLG KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC
720	2070	A .	5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN
722	2072		5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
40,,00			,	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	Sequence	/=possible nucleotide deletion, \=possible
1	ļ			sequence		nucleotide insertion
				scidnence		
1	1	İ			ĺ	QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL
	ĺ	ĺ				GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
į						AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
İ				ŀ		PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN
i				į	İ	KAEINPRTNGDRAACQTRTLRFVSNVTEADRI
1					ļ	LLRWERYEPLEARDLLSFIVYYKESPFQNATE
1				i		HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
1	i				i	ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS-
						PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
						KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
i						YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD
						CCPCOHPPPGQVLPPLEAQEASFQKKFENFLH
			1		· .	NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL
				l		RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY
						RIDIHACNHAAHTVGCSAATFVFARTMPHRE
				l .		ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
1					l	ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
						HLALLPPGNYSARVRATSLAGNGSWTDSVAF
	ŀ					YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
					ŧ	LGFFYGKKRNRTLYASVNPEYFSASDMYVPD
]	EWEVPREQISIRELGQGSFGMVYEGLARGLE
						AGEESTPVALKTVNELASPRECIEFLKEASVM
						KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
1						GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
] ,						AGEIADGMAYLAANKFVHRDLAARNCMVSQ
						DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
1						VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV
						TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
i						PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL
						RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
j l						TPRDCSPQNGGPGH
723	2073	Α	5672	1	216	LAWIDNILPEKEKKETDKKRKKKKGAHEDCD
						EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR
i						SSSWSYSL
724	2074	Α	5704	4235	940	ARGRRSRPVWAASWGGRGRPAARRRPRGLA
						ATMGFELDRFDGDVDPDLKCALCHKVLEDP
] :						LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
						LSAKELNHVLPLKRLILKLDIKCAYATRGCGR
						VVKLQQLPEHLERCDFAPARCRHAGCGQVLL
						RRDVEAHMRDACDARPVGRCQEGCGLPLTH
						GEORAGGHCCARALRAHNGALQARLGALHK
						ALKKEALRAGKREKSLVAQLAAAQLELQMT
						ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE
1						
						ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG
						SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN
						GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT
					•	PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
j						KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
1						GDIHQEMDREELELEEVDLYRMNSQDKLGLT
						VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG
					-	DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
1					l	ARAELQLDEGWMDDDRNDFLDDLHMDMLE
				!		EQHHQAMQFTASVLQQKKHDEDGGTTDTAT
1						ILSNQHEKDSGVGRTDESTRNDESSEQENNG
						DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
				İ		NESFISADCTDADYLGIPVDECERFRELLELK
						COVKSATPYGLYYPSGPLDAGKSDPESVDKE
						LELLNEELRSIELECLSIVRAHKMQQLKEQYR
1				1		ESWMLHNSGFRNYNTSIDVRRHELSDITELPE
1						KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR
Щ	L				L	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AAEGISCPSSEGAVGTTEAYGPASKNLLSITE DPEVGTPTYSPSLKELDPNQPLESKERRASDG
						SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR DRLLRERALKIREERSGMTTDDDAVSEMKM GRYWSKEERKQHLVKAKEQRRRREFMMQSR LDCLKEQQAADDRKEMNILELSHKKMMKKR NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF LSVTTV
725	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP WNTDRCFSNYSMVNTTNMTSAVVEFWERN MHQMTDGLDKPGQIRWPLAITLAIAWILVYF CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI
726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PONTFLGTIIRKFEGONKKFIIANARVONCAII YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	A	5716	3	274	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE
728	2078	A	5737	1899	649	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD
729	2079	Α	5741	1 .	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGÄPLSARAAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

ARAAPAALCAFRFADVRAARARARTACIPE APDVVAVLDSVQGTGAPCREKLNIQLOPEQ LDCGAAHLQHPISLQPLKATTVFRAARGLTSV AVASVNNYTAVGETVHRHUKINESMQ VVSRRVVTVAYGEPVHHVMQFDPADSCYLV LMTSHQMARVAVAACNVHSTCODCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGVAS SEGPSRCPAMTVLPSEIDVRQETPGMINIGVAS LPSISGMEMAGLOYGNNRIVTARVPGPAPGHG LAVCNLLERDQFPPFPRQDHVTVEMSVRY GRNIVKANFTIYDCSSTAQVYPHACTSCLS, QWPCFWCSQQHSCVSNQSRCEASPNTTSDD CPKTLLSPLAPVTGGSQNILLYARVPGPAPGHG LAVCNLLERDQFPFPFPNQDHVTVEMSVRY GRNIVKANFTIYDCSSTAQVYPHACTSCLS, QWPCFWCSQQHSCVSNQSRCEASPNTTSDD CPKTLLSPLAPVTGGSQNILLYARVPGPAPGHG AALECSFGLEEIFEAVWVMESVVRCDQVVIL TIKKSQVPFISLQLKGRPARTLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDG RLRGPLQPMAGTCPAPEIRAETPSPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPP DRYTYSEEIVCVTOPAPGPLSQVTVTVASKE GKSRDRFSVVLPLHSLEFTHMPFKAGGTRITH HGNDLHVGSELQVLNDTDTDFTEMBTROTSI ACTMPEGALPAVPVCVRFERRGCVHGNIT HGNDLHVGSELQVLNDTDTDFTLMKTDTSI ACTMPEGALPAVPVCVRFERRGCVHGNIT WYMQNYSMAVHLIGREPTLCXVLNSTLITCH WYMQNYSMAVHLIGREPTLTCXVLNSTLITCH WYMQNYSMAVHLIGREPTLCXVLNSTLITCH WYMQNYSMAVHLIGREPTLACXVLNSTLITCH WYMQNYSMAVHLIGREPTLACXVLNSTLITCH WYMQNYSMAVH			_				
nucleoide sequence oil de sequence of the peptide correspondia plate of the peptide sequence of peptide se							• • •
USSN (1946) 914 1916 1917 1918 10-cation go of first atmine said of residue of peptide sequence peptide			hod	ID NO:			
sequence 09/496 corresponding on the sequence peptide residue of peptide sequence pept	nucl-	peptide	,		nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence 914 mg to first amino each residue of peptide sequence Pe	cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
minto seid residue of peptide sequence piptide sequence p	seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
mino seid residue of peptide sequence plants of the sequence of peptide sequence plants of the sequence plants of the sequence sequence plants of the sequence p	uence	-	l	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
residue of peptide sequence P-Tyrosine, X-Unknown, "Sop odoon, "possible nucloside delicin, "possible nucloside delicin, "possible nucloside delicin, "possible nucloside delicin, "possible nucloside delicin, "possible nucloside delicin, "possible peptide insertion FDIAPSDDNIKLIKQGAKEGPIKLGFVSAFLHI SIPPPERQUSYAYLALNISEARAGDKESQARSL LARICLPHGAGGDAKKITESVIQLGLQGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG			i			of peptide	
peptide	ĺ		ĺ	[
nucleotide insertion FDINIPSDIDNIA IKINGGAKEOHKL GFVSAFLH SDPPPGAQSY AYLALINSEARAGINESQARSL LARICLPHGAGGDAKKLTESYUQL GLCAGG AGRODLYSRLVSVPPARERLPAVPERPOGSPY ARAPAAL CAPERADVRAHRARATACFVER APDVVAVILDSVVQGTGPACERKLINIOLOPEE LDCGAAHLOHFLSILQPLKATPYPRAPGLTSY AVASVINIYTAVFLGTVINGELLKININISMQ VVSRRVYTVAYGEPVHHVMQFDFADSGYTL LMTSRIMGMARKVAACNVBSTCGOLCYGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTTSA SGGPSRCPAMTVLLSFEDWRQEFYGRMLQISGS LPSLSGMEMACDYGNNIRITVARVPGPAFGIKSY AVASVINIYTAVFLGTVINGSQHFWTTSA SGGPSRCPAMTVLLSFEDWRQEFYGRMLQISGS LPSLSGMEMACDYGNNIRITVARVPGPAFGIKSY AVASVINIYANFTUDCSRTAQVYPHTACTISCLSS QWPCFWCSQQHSCVSNOSRCEASPHTSTPD CPRTILLSFLAPVPTGGSQNILVFLANTAFFCGG AALECSFGLEEFFEARVVNIESVYRCDQVVLH TTRKSQVFPLSLQLKGRPARFLDSFPEMTVW VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPERRAPLSSPDLOGGT LLTIRGRANGRISLSVANVNIESVYRCHGWYSEDY TYNSEEIVCVTGPAFGLASVTVVNASKE GKSRDRFSYVLPLVYSLSEPTMGFKAGGGRITT HGNDLHVGSELQVLVNDTIDPCTELMRITDTSI ACTMPEGALPAPVPVVVSERRRGCVGGNSTLTVAGERFH MVQNVSANVHHIGREPTLCKVLNSTLTICPSI GALSNASAPVDFFRIGRAYADEVAVAEELLD PERAGRGSRFRLDVPVVSFERRGCVGGNTTVAGSER WYMQDVTTAISPRRSSVSGGRITTVAGERFH MVQNVSANVHHIGREPTLCKVLNSTLTICPSI GALSNASAPVDFFRIGRAYADEVAVAEELLD PERAGRGSRFRLDVPPVQFSTRKREWIKH HPGEPLTLVHVSTKGAGKGDSLGLQSHEY RVKIGQVSCDIQVSDRJIHIGSVNSELGAAVG LPITIQVGNFNOTHATLQLGGSSTAINVSIVICSS LLLLSVVALFVFCTKSRRAERVWGRTLLQMR GSGSGAGGSTGDFLEVKHPVTTFFYRCSSLCERETVLPSOT LNSGSSGAGGTHPLLGEWKJPSCGRNBME GISLFSSLLDNKFHLLVFVILLADDRTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKNDLDTESTVLNSPGG MDSLSVRAMDIDTLTQVKKKLEAFCKNVPY SQWPRAEDVOLEWRASTSTONLTRDLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKNDLDTESTVLNSPGGG MDSLSVRAMDIDTLTQVKKKLEAFCKNVPY SQWPRAEDVOLEWRASTSTONLTRDLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKNDLDTESTVLNSPGGG MDSLSVRAMDIDTLTQVKKKLEAFCKNVPY SQWPRAEDVOLEWRASTSTONLTRDLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKNDLDTESTVLNSPGGGG MDSLSVRAMDIDTLTQVKKLEAFCKNVPY SQWPRAEDVOLEWRASTSTONLTRDLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKRVLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKRVLDDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKXD NTLGRVKRVLDDTTSV VE					3	Sequence	
FDLNPSDDNILKIKQGAKEOHKI GFYSAFIH SPPFPGQSYAYALNISABARGARSSAQASSI LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRILVSYPARRIFLASPERPGGSPA ARAPAALCAFRFADVRAAIRAARTACFVEP APDVVAVLDSVVQGTGPACEKKILNIGLQPE APDVVAVLDSVVQGTGPACEKKILNISMQ VVSRRVYTVAVGEPVHHVMQFDPADSGYLY LAYSAYNNYTAVFLGTVAGRLLIKINISMQ VVSRRVYTVAVGEPVHHVMQFDPADSGYLY LAYSAYNNYTAVFLGTVAGRLLIKINISMQ VVSRRVYTVAVGEPVHHVMQFDPADSGYLY LAMISHOMAPKVAVACNYBYGGGCVGAAD AYGGWCALETRCTLQQDCTNSSQOHFWTSA SEGPSRCPAMTVLPSEDDVRQSYPGMLISIGS IPSLSGMEMACDYGNNIKTVARVPGFASCHE IAYCALLPRDQFPPPPPDNDHVTVEMSVRVN GRNIVKANFTTVDCSRTAQVYPHTACTSCLSA GUPPETULSPLAPPPTQGSQNILVPLANTAFFGG CPRTILSPLAPPPTQGSQNILVPLANTAFFGG CPRTILSPLAPPPTQGSQNILVPLANTAFFGG CPRTILSPLAPPPTQGSQNILVPLANTAFFGG ALLESQVFPLSLQLKGRPASFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCMWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSEPMTVM VYYACAMGSPDCSQLLGFEDLGHLCAWSDG RRGGPLSPMAGTCFAPFBRAFLDSTDCT LLTRGNILGRALDSAFLDSTDCTATAGTCFAPFBRAFLACTAGTCFAPFBRAFLATATAGTCFAPFBRAFLATAGTCFAPFBRAFLATAGTCFAPFBRAFLATAGTCFAFTAGTCF		!	1		;	ŀ	
SDPPFGAQSY AYLALNISEARAGDKESQARSL LARICLPHGAGGBAKLTESYIQLGLQCAGG AGRODLYSRLVSVPPARERLFAVPERPQGSP ARAPAALCARFEADVRAARRAARTACEVEP APDVVAVLDSVVQGTGPACERKLNIQLQPEG LDCGAAHLQHPLSLQPLKATPVPFAGGLTSV AVASVNINYTAVELGTVINGELLKINLNESMQ VVSRRVVTLAVGEGVHHVMGEPDABGYLY LMTSHOMARVKVAACNVHSTCGDCVGAAD AYCOWCALETRCTLQOCTINSSQOFHWTSA SEGPSRCPAMTVLPSEDVRQEYPGMILQISGG LPSLSGMEMACDYGNIRITVARVPGPAGGHT AYCOLLPROPEPPPRODHVTVEMSVRVN GRNIVKANFTTUDCSTRAQVYPHTACTISCLSS QWPCFWCQQGISCVSNQSRCCASPSINTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFPCGG AALECSSTLOEUERAVVNESVNCCLQVLH TTRKSQVFPLSLQLKGRPARFLDSPEMTIVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAEPLSQPLDGGT LTLTRGRBNLGRALSDVAHGVWIGGVACEPLP DRYTVSEBIVCVTGPAGGLASVVNTINTAGERPH MYQNYASSERVYLVSVSVNCACEPLP DRYTVSEBIVCVTGPAGGLASVAVAELLD ACTMPEGALPAPVPVCVFERRGCVGCFGBLT HGRDLHVGSELQVLVNIDTIDFCTELMRTDTSI ACTMPEGALPAPVPVCVFERRGCVGCFGBLT WYMQNPVTTAISPRSSVSGGRTITVAGERPH MYQNYSMAVHHIGBEPTLCKVLNSTLTICFSI GALSINASAPVDFFINGRAYADEVAVAEELLD PEAQORGSRFRLDVPVPQFSTARKEWIKH HPGEPLTLVHVSTKGAGKGDSLGLQSHEY RVKIGQVSCDQIVDSGHIHCSVNESLGAAVG LPTTIQVGNFROTHATLQLGGSSTAIIVSTVICSS GIPFLEVKHFVTRTFPFYCSSLLEERFVLPSG LILLSSVVALFVPCTKSRRAERTWQRTLLQMK GISLFSSLDANKFILLGENERSVENGTLLQMK GISLFSSLDANKFILLGENERSVENGTLLQMKSDGAFAVB GISLFSSLDANKFILLGENERSVENGTLLQMK VSCLETYGEPFFILLCALKQQINKGSIDAITG KARYTLNEEWLIJRENGARSTNANVSPGGGG MDSLSVRAMDTDTLTQVKEKILEAFCNAVPY SQWPRAEDVDLEWWASSTGSVILRQNEDDDTSV VEGGRKKLNTLAHYKIPGGSLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLLSTKGTLQKFLDDTSV VEGGRKKLNTLAHYKIPGGASLAMSLIDKKD NTLGRKKVLPSPTLTRLATKGTLQCHEQVALM RCMAHLAESSRYQNGENTINVAMAEENKYRK NTRDJINALAESSRYQNGENTINVAMAEENKYRK NTRDJINALAESSRYQNGENTINVAMAEENKYRK NTRDJINALAESSRYQNGENTINVAMAEENKYRK NTRDJINALAESSRYQNGENTINVAMAEENKYRK NTRDJINALAESSRYCQLGCHTGT			<u> </u>	<u></u>	sequence		
LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGROLYSRILVSYPARRICHAPVERPAGSPY ARAAPAJCARPRADVRAARAARTACFVEP APDVVAVLDSVVQGTOPACEKKINIGLQPE LDCGAAHLQHEJSLQPLKATPVFRAPGLTSY AVASVNYTAVELTVTAVGELIKINISMQ VVGRRVYTVAVGEEVHHVMQFEPADSGYTYL LMTSHQMARVKVAACNVHSTCGDCVGAAL AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGFSRCPANTVLFSEIDVRGEYFGMILQISGS LPSLSGMEMACDVGNIKTVARPVGPAFGIK AYCHLPROPFPFPNQDHTVLEMSVRVN GRNIVKANFTTUDESTRQVYPHTACTSCLSS QWPCFWCSQQHSCVSNQSRCBASPNTTSPQD CPRTILSPLAPPYTGGSQNIKUVLANFOGFAFGIK IAYCHLPROPFPFSPNQDHTVLEMSVRVN GRNIVKANFTTUDESTRQVYPHTACTSCLSS QWPCFWCSQQHSCVSNQSRCBASPNTTSPQD CPRTILSPLAPPYTGGSQNIKUVLANFAFGG AALECSFGLEEFFAVWVNSSVVRCDQVVLH TTRKSQVPPLSQLKGRPAFLDSPFPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLGGFLQPMAGTCPAFERAFEPLSPLDGGT LLTIRGRNLGRRLSDVAHGWWIGGVACEPLP DRYTYSEEIVCVTGAPERAFEPLSPLDGGT LLTIRGRNLGRRLSDVAHGWWIGGVACEPLP DRYTYSEEIVCVTGAPERAFEPLSPLDGGT LLTIRGRNLGRRLSDVAHGWWIGGVACEPLP DRYTYSEEIVCVTGAPERPAGLSPLDLGGT HONDOWNSMAVHHIGREPTLCKVLNSTLTICTSS GALSNASAPVDFFRIGRAYADEVAVEELLD PEBAGRGSRRLDYLPNPOPSTAKREKWIKH HPOEPLITUHVNYTKAGRSCDNSLGSHEY WYNGNASAPVDFFRIGRAYADEVAVEELLD PEBAGRGSRRLDYLPNPOPSTAKREKWIKH HPOEPLITUHVNYTKAGRSCDNSLGSHEY RVKIGQVSCDIQVSDRIHCSVNESLGAAVGC LPTITQVGNPRQTATTDQGRCRTANSVTCSS GIPPLEYKHFYTRTPFCSSLLEERYVLNSQT LNSQGSSQAGGTHPLLGEWKIPSCRRNMER GISLFSSSLLDNKHFLLVEVALLEQQGCRFAVKD RCSLASLLTIALHGKLEYYTSIMKELLYDLID ASAAKNPLLMRRTSVVEKMLTNWSTGGG MDSLSVRAMDTDTLTQVKEKLEACKNVPY SQWPRAEDVDLEWPASSTGSVILRDLDDTSV VEGGRKKLNTLAHYKIPEGSLAMSLIDKKD NTLGRWRDLTSVAGREGNERSTGNANSVGGGG MDSLSVRAMDTDTLTQVKEKLEACKNVPY SQWPRAEDVDLEWPASSTGSVILRDLDDTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRWRDLTSVRYAKGRENTWALAGERYRGNLDDTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRWRDDTTLSTSKERTENATRATCLQCHKGFLODD- KARYTLNEEWLLBENBARRRINNYSGGGG MDSLSVRAMDTDTLTQVKEKLEACKNVPY SQWPRAEDVDLEWPASSTGSVILRGLDDTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRWRDDTTLSTSKERTENATRATCLQCHKGFLORVAL MEDNITICCTSGA A 5744 3 292 QFPHFFISHELTGLLRTAQLPEQVSWPWGQ VANGKGNQSNMGSPQSLLAFERNLELQDNG MAHLALEESRCYQDETNTWALAGERYRINN AHELAESRCYCHAPTTNAALAGERYRING NAHLAEESRCYGDETNTWALAGERYRING VANGENGENDETNTWA	1		i	i .		1	
AGRODLYSRLVSVPPARERLJAVPERPÖGSP ARAPAALCAPRADVRANTARATACEVEP APDVVAVLDSVVQGTÖPACERKI.NIGLÖPEE LDCGAAH.OHPI.SLQ.DPLATPYPERAGI.TSV AVASVNNYTAVFLGTVNGRLI.KINLINESMQ VVSRRVYTVAYGEPVHHVMQFDPADSÖTYL. LMTSHQMAADVVAACNIVHSTCGDCVGAAD AYGGWCALETRCTLQQDCTNSSQQHFWTTSA SEGPSRCPAMTVLPSEDDVRQEYPGMLQISGS LPSLSGMEMACDYGNNIRITVARVPGPAFGIR IAYCOLLPRDOPPPPPPNODHVTVEMSVRVN GRNIVKANFTIVDCSRTAQVYPHTACTISCLS, QWPCFWCSQQHSCVSNOSRCBASPNTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFGGG AALECSFOLEEFERAVVNESVYRCDQVVLH TTRKSQVPPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLGMWSDGG RIRGPLOPMAGTCPAPERAEPLSCPPUTVM VYNCAMGSPDCSQCLGREDLGHLGMWSDGG RIRGPLOPMAGTCPAPERAEPLSCPLOGVTH HONDLHVOSSELQVLVNDTDPCTEHARITOTSI ACTMPEGGALFAPPVCVRFBRRGCVFGSLTT WYNMONPVTIAISPRSYVSGGRITVAGEREP MYQNVSMAVHHIGREPTLCKVLNSTLTICPS GALSNASAPVDFFRINGRAYADEVAVAEELLD PEEAQRGSRRRLDYLPNPQFSTAKREEWIKH HPGEPLTLVHIVSTKAGAGKGODSLGJSHEY RVKJGQVSCDIQIVSDRIHICSVNESLGAAVGG LPTITQVGFFRYATHALQGGSETAINISVICSS LLLLSVVALFVFCTKSRRAERYWQRTLJQUMB EMESQUREERIKGFBALQTDMTDLTKELNRSC GIPFLEVKHFVTRTFFPKCSSL/ERRYVJLPSQI LNSQGSSQAGPTHPLIGWKHESEGLGSHEY RVKJGQVSCDIQIVSDRIHICSVNESLGAAVGG GISLFSSLLDNKHFLIVFVHALDQGCDFAVED GCSLASLLTTALHGKLEYTSINKELLJVLDLD ASAAKNPKLMLRTTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIRQQINKGSDATIG KARYTILNEWLLABENGARFNINNYSTQGGG MDSLSVRAMDTDTI.TQVKEKLEARCKNIVYP SQWPRAEDVDLEWFRASTGSYLIRGLDDTSV VEDGRKKLNTLAHYKPPGGASLAMSLIDKKD NTLGRVKDDLTEETYFHLVFTPLALEFRKSH RQSHRKKVLEPIYLTRLLSTKGTLQKFLDDLTSV VEDGRKKLNTLAHYKPPGGASLAMSLIDKKD NTLGRVKDDLTEETYFHLVFTPLALEPRKSH RQSHRKKVLEPIYLTRLLSTKGTLQKFLDDLTSV VEDGRKKLNTLAHYKPPGGASLAMSLIDKKD NTLGRVKDDLTEETYFHLVFTPLALEPRKSH RQSHRKKVLEPIYLTRLLSTKGTLQKFLDDLTSV VEDGRKKLNTLAHYKPPGGASLAMSLIDKKD NTLGRVKDDLTEETYFHLVFTPLALEPRKSH RQSHRKKVLEPIYLTRLLSTKGTLQKFLDDLTSV VEDGRKKLNTLAHYKPPGGASLAMSLIDKKD NTHINGTNASLPLRFWNILKNPQFYFDLICAGARGGE PDTLHHWKTNSLPLRFWNILKNPQFYFDLICAGARGGE PDTLHHWKTNSLPLRFWNILKNPQFYFDLICAGARGGE PDTLHHWKTNSLPLRFWNILKNPQFYFDLICAGARGGE PDTLHHWKTNSLPLRFWNILKNPQFYFDLICAGARGGE PLALEFRKSH RQSHRKKVLPRFTNTMAAGARGRYKYAK RYPPOMAALLAAPPTARRTQLQHKFEYVAL RENDTYCYSEA PDTLHHWAT			l			İ	
ARAAPAALCAFRFADVRAARARARTACIPE APDVVAVLDSVQGTGAPCREKLNIQLOPEQ LDCGAAHLQHPISLQPLKATTVFRAARGLTSV AVASVNNYTAVGETVHRHUKINESMQ VVSRRVVTVAYGEPVHHVMQFDPADSCYLV LMTSHQMARVAVAACNVHSTCODCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGVAS SEGPSRCPAMTVLPSEIDVRQETPGMINIGVAS LPSISGMEMAGLOYGNNRIVTARVPGPAPGHG LAVCNLLERDQFPPFPRQDHVTVEMSVRY GRNIVKANFTIYDCSSTAQVYPHACTSCLS, QWPCFWCSQQHSCVSNQSRCEASPNTTSDD CPKTLLSPLAPVTGGSQNILLYARVPGPAPGHG LAVCNLLERDQFPFPFPNQDHVTVEMSVRY GRNIVKANFTIYDCSSTAQVYPHACTSCLS, QWPCFWCSQQHSCVSNQSRCEASPNTTSDD CPKTLLSPLAPVTGGSQNILLYARVPGPAPGHG AALECSFGLEEIFEAVWVMESVVRCDQVVIL TIKKSQVPFISLQLKGRPARTLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDG RLRGPLQPMAGTCPAPEIRAETPSPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPP DRYTYSEEIVCVTOPAPGPLSQVTVTVASKE GKSRDRFSVVLPLHSLEFTHMPFKAGGTRITH HGNDLHVGSELQVLNDTDTDFTEMBTROTSI ACTMPEGALPAVPVCVRFERRGCVHGNIT HGNDLHVGSELQVLNDTDTDFTLMKTDTSI ACTMPEGALPAVPVCVRFERRGCVHGNIT WYMQNYSMAVHLIGREPTLCXVLNSTLITCH WYMQNYSMAVHLIGREPTLTCXVLNSTLITCH WYMQNYSMAVHLIGREPTLCXVLNSTLITCH WYMQNYSMAVHLIGREPTLACXVLNSTLITCH WYMQNYSMAVHLIGREPTLACXVLNSTLITCH WYMQNYSMAVH			l				LARICLPHGAGGDAKKLTESYIQLGLQCAGG
APDVVAVLDSVVQGTGPACERKINIQLOPAGE LDCCAAHLQHPISLIQPIKATPRAGITSV AVASVNNYTAVELGTVNGRELKINILNENG LDCCAAHLQHPISLIQPIKATPRAGITSV AVASVNNYTAVELGTVNGRELKINILNENG VVSRRVVTTAVELGTVNGRELKINILNENG VVSRRVVTTAVELGTVNGRELKINILNENG VVSRRVVTTAVELGTVNGRELKINILNENG VVSRRVVTTAVELGTVNGRELHKINILNENG AVACOWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQFTSGMILGISG LPSLSOMEMACDYGNNIRTVARPGPAGE LPSLSOMEMACDYGNNIRTVARPGPAGE AVAILABOUT SERVEN GRNIVKANFTTYDCSETAQVYPHTACTSCL QWPCPWCSQQHSCVSNGSRCSPRSPTSPQD CPRTLLSPLAPVTTGGSQNILVPLANTAFTCG AALECSFGLEEIFEAWWVSESVRCDQVVLH TTRKSQVPFLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSCCLGREDLGHMSDGG RLRGPLQPMAGTCPAPERAGEPLSQPLOGED LLTIRGRNLGRRISDWAHGWWGAVCEPLP DRYTVSEEIVCVTQPAPGELSQVVTVMSSEE GKSRDRSSVLPHVHSLEETMFCAMSDGG RLRGPLQPMAGTCPAPERAGEPLSQPLOGED LLTIRGRNLGRRISDWAHGWWGAVCEPLP DRYTVSEEIVCVTQPAPGELSQVVTVMSSEE GKSRDRSSVLPHVHSLEETMFCAGGGRITTWAGGRHI MVQNVSMAVHHIGREPTLCKVLNSTLITGST GALSMASAFVDFTRNGRAYADEAGTVAFUCARERGCVHGAVIT WYNCAMGSPRSPSVSGGTTTVAGGRHI MVQNVSMAVHHIGREPTLCKVLNSTLITGST GALSMASAFVDFTNGRAYADEAVAVAEELLD PEEAAGGSRFRLDVLPNPQFSTAKREWIKH HPGEPLTLVHIVSTKGAGGCGDSCLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVG LDTTUNGGSSQAGETFINGRAYADEAVAVAEELLD PEEAAGGSRFRLDVLPNPQFSTAKREWIKH HPGEPLTLVHIVSTKGAGGCGDGSTAKREWIKH HPGEPLTLVHTSTKGAGGCAGATANISTUNGS LLLLSVVALFYFCTKSBRAERYWQKTLLQWBP RVSLAGFTRUTTTTTFPCKSSLYEERYYLPSGJ LNSGSSQAGETHLLGGEWRFSKERIKWELLVDLD ASAAKVPKLMLRRTESVVEKMLTNWMSICM YSCLAETHLGGEWRFSKELLVDLD ASAAKVPKLMLRRTESVVEKMLTNWMSICM YSCLAETHLGGWERGRKGGNAMSCHWKGLDVS VEDGRKKLNTLAHTKIEFGASLAMSLIDKW LDVDLD ASAAKVPKLMLRRTESVVEKMLTNWMSICM YSCLAETHLGGWERGRKGGNAMSCHWKGLDVSVEW WGQ WANGKGAFBACSISDLQLGKDSPINK RLLYAKEPEYRTVLPSGJLAFEKKSH RQSHRKKVLPETVLTRLLSTKGTLQKFCDSDINK TDHIDACLSVLAQAFBACSISDLQLGKDSPINK RLLYAKEPEYRTVLAVETTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT						ĺ	AGRGDLYSRLVSVFPARERLFAVFERPQGSPA
LDCGAAHLQHPISILQPIKATTYPRAFQIS AVASVNNYTAVGETVHNYLKININESMQ VVSRRVYTVAYGEPVHHVMQFDPADSGYL LMTSHQMARVKVAACNVHSTCGQCYGAAD AYCGWCALETRCTLQQDCTNSGQCHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGLPSISGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTMINIGYSGUPFTMINIGYSGUPFTMINIGHTHATCHSCLSA QWPCFWCSQQHSCVSNQSRCEASPPHTSPGD CPRTLISPLAPVTGSSNIGYSTUPLANTFTPGG AALECSFGLEEIFEAVWWESVVRCDQVVLH TTRKSQVPFLSLQLKGRAPTLATGTSG ALECSFGLEEIFEAVWWESVVRCDQVVLH TTRKSQVPFLSLQLKGRAPTLATGTSGUPFTMINIGYSGUPFTMINIGYSGUPTTMINIGTPGMINIGTPTLATGTSGUPFTMINIGYSGUPTMINIGYSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTTLATGT							ARAAPAALCAFRFADVRAAIRAARTACFVEP
LDCGAAHLQHPISILQPIKATTYPRAFQIS AVASVNNYTAVGETVHNYLKININESMQ VVSRRVYTVAYGEPVHHVMQFDPADSGYL LMTSHQMARVKVAACNVHSTCGQCYGAAD AYCGWCALETRCTLQQDCTNSGQCHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGLPSISGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGQHFWTSA SEGPSRCPAMTVLPSEIDVRQETPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTPGMINIGYSGUPFTMINIGYSGUPFTMINIGYSGUPFTMINIGHTHATCHSCLSA QWPCFWCSQQHSCVSNQSRCEASPPHTSPGD CPRTLISPLAPVTGSSNIGYSTUPLANTFTPGG AALECSFGLEEIFEAVWWESVVRCDQVVLH TTRKSQVPFLSLQLKGRAPTLATGTSG ALECSFGLEEIFEAVWWESVVRCDQVVLH TTRKSQVPFLSLQLKGRAPTLATGTSGUPFTMINIGYSGUPFTMINIGYSGUPTTMINIGTPGMINIGTPTLATGTSGUPFTMINIGYSGUPTMINIGYSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTPTLATGTSGUPTMINIGTTLATGT			l			Ì	APDVVAVLDSVVOGTGPACERKLNIOLOPEO
AVASVNNYTAVELGTVNGRLIKININESMQ VVSRRVYTAVELGYHVHMQPDAGSYLY LMTSHQMARVXVAACNYHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQPTYGMILQISG LPSLSGMEMACDYGNNIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK LAYCNLLPROJPPPPNQDMIRTVARVPGPAGIK GRINVKANFTIYDCSRTAQVYPHTACTSCLS QWPCFWCSQQJBCVSNGSRCGSSPNTTSPQD CPRTLISHLAPVTIGGSGNILVPLANTAFFQG AALECSFGLEEIPEA WWNESVCDQVULH TTRKSQYPPLSLQLKGERAFLDSPEPNTYM VYNCAMGSPDCSQCLGREDLGHLCMWSDQ RIRGPLQPMAGTCRAPERAGEHSGPLDGGT LLTIRGRRLGRRLSDVAHGVWIGGVACEPP DRYTYCSERIVCYTOPAPCHSGTVTVANASKE GKSRDRFSSYVLPLVHSLETTHOPKAGGTRITT HGNDLHYGSELQULVNDTDCTTELMRTDTSI ACTMPEGALPAVPVCVXFERRGCVHGNLT WYMQNYMAVHHIGREPTLCXVLNSTLITCHT WYMQNSTALPAPVPVCVXFERRGCVHGNLT WYMQNSMAVHHIGREPTLCXVLNSTLITCHT WYMQNSMAVHHIGREPTLCXVLNSTLITCHT WYMQNSTALPAPVPVCVXFERRGCVHGNLT WYMGNSTALPAPVPVCVXFERRGCHGSTLITVAGERF WQNNSMAVHHIGREPTLCXVLNSTLITCHT WYMGNSTALPAPVPVCVXFERRGCHGSUSLGLAVG LPTIQVGNSDQLIVSDRIHCSVXFWWQKTLLQME EMESQIREERKGFAELQTMDTLITKELNRSC GIFFLEYKHFVTTFFFYCKSSLJCAAVG LPTIQVGNSDQLITCHTURGSTSCLAAVG LPTIQVGNSDQLITCHTURGSTSCLAAVG LPTIQVGNSDQLITCHTURGSTSCLAAVG LSVALMPTTTFFFYCKSSLYGWQKTLLQME EMESQIREERKGFAELQTMDTLITKELNRSC GIFFLEYKHFVTTFFFYCKSSLYGWQKTLLQME EMESGIREERKGFAELQTMDTLITKELNRSC GIFFLEYKHFVTTFFFYCKSSLYGWQKTLLQME EMESGIREERKGFAELQTMDTLITKELNRSC GIFFLEYKHFVTTHFFYFYCKSSLYGWQKTLLQME GISLFSSLLDNKHTLIVFVIALLGCKGVAPVPD LNSQGSSQAQETHPLLGGWKPESCRPNMEE GISLFSSLLDNKHTLIVFVIALLGCKGVAPVPD LNSQGSSQAQETHPLLGEWKPESCRPNMEE GISLFSSLLDNKHTLIVFVIALLGCKGAPVPV SQWFRAEDVDLEWFASSTQSYLLRDLDDTST VEGGGRKLIVTLAHTKJPFPCKSSLLAWCHNIVSTOGG MDSLSVRAMDTDTLTQVKERLLAFCKNVPY SQWFRAEDVDLEWFASSTQSYLLRDLDDTST VEGGGRKLIVTLAHTKJPFPCKSSLLAWCHNIVSTOGG MDSLSVRAMDTDTLTQVKERLLAFCKNVPY SQWFRAEDVDLEWFASSTQSYLLRDLDDTST VEGGGRKLIVTLAHTKJPFPCTSDLAFKSH RQSHRKKVLPEPTTTLLTLSTTGTLEAPTATATATATATATATATATATATATATATATATATAT			ļ			ŀ	
VYSRRVYTVAGGEVHHVMQDPADSGYLY LMTSHQMARVKAACNYGDCVGALO AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTYLPSEDVRQFYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFIGH IAYCNLPRDQFPFPPNQDHVTVENSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPPHTSPQD CPRTLISPLAPVPTGGSGNILVPLANTAFFQG AALECSFGLEEFFAVWYNESVVRCDQVLH TTRKSQVFPLSQLUKGRFARLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDG RLRGPLQPMAGTCPAPEIRAFLSGPLDGGT LLTIRGRUGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRSTVALPLVHSLEPTMGFKAGGTRIT HGNDLHVGSELQVLVNDTDPCTELMRITDTSI ACTMPEGALFAPVPVCVSRGCVTGNASKE GKSRDRSTVALPLVHSLEPTMGFKAGGTRIT WYMQNDVITAISPRRSPVSGGRTITVAGERH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPGTSTAKREKWIKH HGEPLTLVHWSTKGAGKEGDSLGJCSHEY RVKIGGVSCDIGVSDRIHGVSVESLGAAVQ LPITIQVGNFNOTIATLQLGGSETAIVSTVICSV LLLLSVVALEVPCTKSRRAERYVQKTLLQMB EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTTRTFFFKCSSLYEERXVLPSQT LNSQGSSQAGETHPLLGGWKIPSCRSPMWEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIAHCKLEFYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLETVGEFFFILLCAIKLAPCKNVYS SQWPRAEDVDLEWFASSTGSYILRDLDDTSV VEDGRKLLNTLAHKVERGASLAMSLDKKD NTLGRVKDLDTEXYFHLVLPTDELEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKPGGVFDIDK TDHDACLSVIAQAFIDACSISDLQLGKDSFTN KLLYAKEIPFLATKLFSVTQRYKQQDMPTISEQE MMAHLAEESRKYQNEFTNINVAMABIYKYAK RYRPGIMAALEANPTARRTQLQHKFEQVVAL MEDNTYSCYSEA 730 2080 A 5744 3 292 QPSPLFISHLETLOLLETAGQLFEQVSWWGQ VANGKGNQRNMGSPPOPSLLAFERNLELQIMC LGYSLLMGKURPRVAKOTILAFRODFPSCV LKD LKD THIDMCLSVIAQAFIDACSISDLQLGKDSFTN KLLYAKEIPFARRTQLQHKFEQVVAL MEDNTYSCYSEA 730 2080 A 5747 I 382 FLECKMRAFRSSKLLQVGYTPDGKDDYRWC			ł	i		ł	, ,
LMTSHOMARVKVAACNVHSTCGDCVGAAD AYCGWCALETRCTI,QQDCTSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISG LPSLSGMEMACDYGNNIRTVARVPGPAFGK IAYCNLLPRDQFPFPPRODHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQGHSCVSNOSRCEASPNFTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEFFAVWVTSVCRVCDVVLH TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTYSEBLVCVTGPAPGPLSGVTTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVOSELQVLVNDTIDPCTELMRITDTSI ACTMPEGALPAPPVCVRFERGCVHGNLTI WYMQNPVITAISPRSPVSGEGTVTVAGGRFH MYQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYBVAYAEELLD PEBAQRGSRFRLDYLPNPOPSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHESVVESLGAAVGG LPTITQVGNFPNQTIATLQLGGSETAIIVSVICSS LLLLSVVALFVCTKSSRAEVAYASELLD EMESQIREERKGFAELQTIDMTDLTKELNSC GIFFLEYKHFVTCFTKSFRAESYLVSTSINGSL LNSGGSSQAGETHPLLGEWKIPESCEPNMEE GISLFSSLLDMKHFLLVFVHALQQKDFAVRDE LNSGGSSQAGGTHPLLGEWKIPESCEPNMEE GISLFSSLLDMKHFLLVFVHALQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLD ASAAKNPKLMLRRTESVVEKMLTNVMSIGM YSCLRETVGEPFFILLGAKQRNKGSIDATIG KARYTLNEEWLLREINEAKRRILAPCKNVTY SQWPRAEDVDLEWFASSTGSYILRDLDDTSV VEDGRKKLNTLAHVKPEGLAFCKNVTY SQWPRAEDVDLEWFASSTGSYILRDLDDTSV VEDGRKKLNTLAHVKPRGSIDALIG KARYTLNEEWLLREINEAKRRILAPCKNVTY SQWPRAEDVDLEWFASSTGSYLRDLDDTSV VEDGRKKLNTLAHVKPRGSIDALIG KARYTLNEWLLREINEAKRRILAPCKNVTY SQWPRAEDVDLEWFASSTGSYLRDLDDTSV VEDGRKKLNTLAHVKPRGSIDALIG KARYTLNEWLREPHYTRLSTYRCLGFKKSH NTLGRVKDLDTEXYFHILVPTDELEAGEKGISD PDTLHWKTNSLPLFFWNAILKNPGVFVDIDK TDHDDACLSVIAQAFIDACSSDLQLGKDSPTN KLLYAKEIPEYNRIVQRYYRQLQDMPTPSGG MMAHLAEESRKTYQNEPTNNVMAEIYKYAK RYRPGIMAALEANPTARTQLQHKFEQVVAL MEDNTYSCYSEA 730 2080 A 5744 3 292 OPSSLFHSHLETLQLLRTAQLPEQVSWWGG VANGKGNQRNMGSPOPSLLAFERNLELQMC LGYSLLMGKLRPRVAKDTLARTGSTSPLI- LKD TJLLGKCLTRAVARDTARTQLJCHGROSTSPLI- LKD LKDCMTRCTSSEA			}				·
AYGGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEDWCFYPGMILQISG LPSLSGMEMACDYGNNIRTVARVPGPAFGH IAYCNLPRDQFPFPPNQDHVTVESWSVN GRNIVKANFTYDCSRTAQVYPHACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPPNTSPQD CPRTLLSPLAPPTGGSGNILVPLANTAFFQG AALECSFGLEEFFAVWNESVYRCDQVLH TTRKSQVPPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRALSDVAHGGVWGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRIT HGNDLHVGSELQVLVNDTDFCTELMRIDTSI ACTMPEGALPAPPVCVERFERGCVHGNLTT WYMQNPVITAISPRRSPVSGGRTTVAGERH MVQNVSMAVHHIGREPTLCKVLNSTLITCFSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNGPSTAKRKWIKH HPGEPLTLVHINSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVG LPITIQVGNFNYGTATLQLGGETAINIVSVICSS LLLLSVVALFVCTKSRRAERYWQKTLLQME EMESQREEBERGFAELQTDIDTIKELARSC GIPFLEYKHFVTRTFFRKCSSLYEERYVLPSQT LNSQGSSQAGEHPLLGEWKIPSCSRFNMEE GISLFSSLLDMKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLD ASAAKNPKLMKLRTESVYKETALLOQDKGSDATIG KARYTLNEEWLLRSPILGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSVGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSPGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSPGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSPGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLLSRINGAKTNIVSPGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLARTHGAKTRTSVKYPGQUPDMPTLSGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVOLEWLARTHGAKTRTSVKYPGQUPDMPTLSGG MDSLSVRAMDTDTLTQVKEKILEAGNA RYRQMAALEANPTARTQQLEQGVWGG MOSLORAMGSGPOPGLAFERNLELQMC LGYSLLMGKLRPRVAKDTLRVHDSTPSPLT LIKD							•
SEGFSRCPAMTVLPSEIDVRQEYPGMILQISG LPSLSGMEMACDYGNMIRLAWPOPAGAGH LAYCHLPRDQPPPPPPNQDHVTVEMSVRVN GRNIVKANFITYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCBASPHYTSPQD CPRTLLSPLAPVPTIGGSQNILVPLANTAFPQG AALECSFGLEEFFAVWVNEXVRCDQVVLH TITRKSQVFPLSLQLKGRPARRLDSPEPMTVM VVNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPERAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRRSTYJLPLVHSLEPFTMGFKAGGTRITI HGNDLHVOSELQVLVNDTDDTCTELMRIDTSI ACTMPEGALPAPVPVCVREERGCVHGNLTF WYMQNSVTHATSPRSFVSGGGTITVAGERFH MYQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYDEVAVAEELLD PEEAQRGSRRLDYLPNPGPSTAKREKWIKH HHGEPLTLVHPNTSTKGAGKDSJCLGVSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVG LPITIQVONFNQTIATLQLGGSETAIVSVICSS LLLLSVVALFVETKSRRAEDVAYAEELLD PEEAQRGSRRLDYLPNPGPSTAKREKWIKH HHGEPLTLVHPYTTRTFPKCSSLYEERVYLPSQT GIFLEYKHFVTCTKSRRAEDVATAGELD EMESQREEIRKGFAELQTIDMTDLTKELNSC GIFPLEYKHFVTTFTFPKCSSLYEERVYLPSQT LNSGGSSQAGGTHPLLGBWKIPESCRPNMEE GISLFSSLLDMKHELLVPHALEQQKDFAVRE RCSLASLLTTALHGKLEYYTSIMKELLVDLID ASAAKNPELMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFILLGRWKIPESCRPNMEE GISLFSSLLDMKHELIVFHALEQQKDFAVRE RCSLASLLTTALHGKLEYYTSIMKELLVDLID ASAAKNPELMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFILLGRWKIPESCRFWMEE GISLFSSLLDMKHELIVFHUNGKLIGAFCKNVYY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHVKPEGALAMSLIDKKD NTLGRVKDLDTEKYFFILLVLPTDELAEPKSH RQSHRKKVLPETYTRILTSKGLLQKFLDDTSV VEDGRKKLNTLAHVKPEGALAMSLIDKKD NTLGRVKDLDTEKYFFILLVLPTDELAEPKSH RQSHRKVLPETYTRILTSKGLLQKFLDDTSV VEDGRKKLNTLAHVKPEGALAMSLIDKKD NTLGRVKDLDTEKYFFILLVLPTDELAEPKSH RQSHRKVLPETYTRILTSKGLLQKFLDDTSV VEDGRKKLNTLAHVKPROSIDQLGCKDSPTM KLLYAKEIDEYSKURVRYQQLQMPTPSGG MMAHLAEESRKYQNEPTNNVMAMEIYKYAK RYRPQIMAALEANPTARTQLQHKFEQVVAL MEDNTYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVWAL MEDNTYECYSEA PGPSLFHSHLETLQLLRTAQLPEQVWAL MEDNTYECYSEA PGPSLFHSHLETLQLLRTAQLPEQVWAL MEDNTYECYSEA PGPSLFHSHLETLQLLRTAQLPEQVWAL MEDNTYECYSEA 11 LKD THE ACCORNICATE STORT TO THE ACCORNICATERSKELLQWGYTPDGKDDYRWCO VANGKGNQRNMGSPQPSLLAFERNLEL QMC TILL LKD TILL LKD TILL LKD TILL LKD TILL LKD TILL LKD TILL LKD TIL						}	•
LPSLSGMEMACDYGNNIRTVANYOPAĞİĞİL IAYCNLLPROQPEPPEPNODHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCBASPNTSPQD CPRTILSPLAPVITGGSQNILVPLANTAFFQG AALECSFGLEEIFEAVWYNESVYNCDQVULH TITRKSQVPPLSQLKGRPAFFLDSFEPMTVN VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RILGPLQPMAGTCPAPERAIPPLSQPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEBIVCVTGPAFGPLSGVYVTNASKE GKSRDRFSYVLPLVHSLEPTMGFKAGGTRIT HONDLHVGSELQVLVNDTIDPCTELMRITDTS ACTMPEGGALPAPVPVVSFERRGCVHGNLTT WYMQNPVITAISPRSPVSGGRTITVAGGRFH MVQNVSNAVHHIGGEPTERGCVHGNLTT WYMQNPVITAISPRSPVSGGRTITVAGGRFH MVQNVSNAVHHIGGEPTERGCVHGNLTT WYMQNPVITAISPRSPVSGGRTITVAGGRFH MVQNVSNAVHHIGGEPTERGCVHGNLTT WYMQNPVITAISPRSPVSGGRTITVAGGRFH MVQNVSNAVHHIGGEPTERGGVHGNLTITOPS GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSSRFRLDVLPNPQFSTAKEKWIKH HPGEPLIL.VHVSTKGAGKEQDSLGJQSHEY RVKIGQVSCDIQIVSDRIHICSVNESLGAAVGG LPTITQVGRNPQTITLVHYSTTGAGKEGDSLGJQSHEY RVKIGQVSCDIQIVSDRIHICSVNESLGAAVGG LPTITQVGRNPQTITLVHYSTTGAGKEGNSUKTLLQWB EMESQRERIKGFABLOTITLTQHSETGALWATINTSVICSV LLLLSVVALFVFCTKSBRABRYWQKTLLQNBG GISFSSLLDINKHFLIVFVHALEQQKDFAVRO RCSLASLLITIALHGKLEYYTISMKELVDLD ASAANNPKLMRRTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQNKGSIDATTG KARYTLNEBWLLRENIESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQNKGSIDATTG KARYTLNEBWLLRENIESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQNKGSIDATTG KARYTLNEBWLLRENIESVYELMELDVDLDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTISVYEKBLLEPKKSH RQSHKKKVLPETVITRLSTKGTLQKFLDDLF KALISREDKPPLAVKYFFDLELEAPKKSH RQSHKKVVLPETVITRLSTKGTLQKFLDDLF KALISREDKPPLAVKYFFDLELEAPKKSH RQSHKKVVLPETVITRLSTKGTLQKFLDGUFFDLYAK KYRPQIMAALEANPTARRTQLQHKFEQVVAL MKDNIYECYSEA MAHALAEESRKVQNEFNINVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA MAHALAEESRKVQNEFNINVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA LKDMYFECYSEA PELKCMRKAFRSSKLLQVGYTPDGKDDYRWC			ļ			ł	1
I AYCNLLPROGPPFPPNODHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTILSPLAPVPTIGGSQNILVPLANTAFPQG AALECSFGLEEIFEAVWNESVVRCDQVVLH TTRKSQVPPLSLQLKGRPARFLDSFPMTVM VYNCAMGSPDCSQCLGEDGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTYSEEIVCVTGPAFGLSGVVTVNASKE GKSRRFSYVLPLVHSLEPTMGFRAGGTRITH HGNDLHVGSELQVLVNDTDPCTELMRIDTSI ACTMPEGALPAPVPVCVFFERRGCVHGNLTF WYYMONPVTAISPRRSPCSGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAFELLD PEFAQRGSRFRLDVLTPMOFSTAKREKWIKH HFGEFLTLVHVSTKGAGKEQDSIGLQSHEY RVKIGQVSCDIQVISDRIHCSVNESLGAAVG LPTIQVGNFNQTIATLQLGGSETAHVSIVICSI LLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQREEIRKGFAELQTDMTDLTKELNSS GIPFLEYKHFVTRTIFFPKCSSLVSERYVLPSQT LNSQGSSQAQETHFLLGEWKIPSSCRPNMEE GISLFSSLLDNKHFLIVFVHALEGQKDFAVRD RCSLASILTIALHGKLEVYTSIMKELLVDLD ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQONKGSIDATIG KARYTLNEEWLLRENBEKRFRNLNVSSPQGG MDSLSVRAMDITDTLTQVKEKILEAFKNVFY SQWPRAEDVDLEWFASGALAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHKKVLPENTITALSTKGTLQKFLDDLF KALLSIREDKPPLAVKYFFDPLEEQAKKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIBK RQSHKKVLPENTITALSTKGTLQKFLDDLF KALLSIREDKPPLAVKYFFDPLEEQAKKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIBK TDHIDACLSVIAQAFIDACSISDLQLGKDSFYN KLLJXAKEREPYRKIVORYKQLOMPTPLSGG MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQLFGCVSWPWGQ ANGKGNRMGSPSPSLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 QFSPLFHISHLETLQLLRTAQLPEQVSWPWGQ VANGKGNRMGSPSPSLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD TILL CKURKAFRSSKLLQVGYTPDGKDDYRWC							
GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTLLSPLAPVPTIGGSQNILVPLANTAFPQG AALECSFGLEEIFEAVWVNESVYCDQVVLH TTRKSQVPPLSQLKGRPAFFLDSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLGWSDGG RLRGPLQPMAGTCPAPERASPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAFGPLSGVYTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVVCVPFERRGCVHGNLTT WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREFTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDVLPPPQFSTAREKWIKH HPGGPLTL.VHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVG LPTTIQVGNFNQTTATLQLGGSETAITVSVICSS LLLLSVVALFVFCTKSRRARTYWQKTLLQME EMESQREEIRKGFABLQTDMTDLTKELNRSC GIPFLEVKHFVTRTFFPKSLYEEKVLPSQT LNSQGSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVVFVALSSLYEERVVLPSQT LNSQGSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVVFVALSCKYEDATAV RCSLASLLTIALHGKLEYYTISIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLGKQNIKGSIBAITG KARYTLNEEWLLRENIBAKPRNLINVSFQGG MDSLSVRAMDTDTLTQVKEKGLKAFCKNVPY SQWPRAEDVDLEWFASSTQSYTLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYTPELAPFKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KARLSTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYTLFULLSTKGTLQKFLDDLF KARLSTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYTLFULLSTKGTLQKFLDDLF KARLSTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYTLFULLSTKGTLQKFLDDLF KARLSTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYTLALSTKGTLQKFLDDLF KARLSTREWNNLENDFPVDIDK NTLGRVKDLDTEKYPHLYTLALSTKGTLQKFLDDLF KARLSTREWNNLENDFPVDIDK NTLGRVKDLDTEKYPHLYTLALSTKGTLQKFLDDLF KARLSTREWNNLENDFPVDIDK NTLGRVKDLDTEKYPHLYTLALSTKGTLQKFLDDLF KARLSTREWNNLENDFPVDIDK NTLGRVKDLDTKKYPHLOGLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKOSPTN KLLYAKEREFYRNICQRYKQQLGKSGRISH CARYTONATATATQLCLGKTAGATATARTQLQLFQVSWPWGQ MANGKGNGMGSSPSPSLLAFFRNLELQIMG LGYSLLMGKLRFRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 FELKCMRAFFSSKLLQVGYTPDGKDDYRWC			l				
QWPCFWCSQQHSCVSNQSRCBASPNFTSPQD CPRTLISPLAPYTGGGSQNILVPLANTAFFQG AALECSFGLEEIFEAVWVNESVVRCDQVVLH TTRKSQVFPLSLQLKGRPAFFLDSSFEMTVW VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRRLGRRLSDVAGWIGWACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRIDTSI ACTMPEGALPAPVPVCVBFERRGCVHGNLTT WYMONPVITAISPRRSPYSGGRITIVAGGERH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELD PEEAQRGSRFRLDYLPHYOPSTAKREKWIKH HHGEPLTLVHNSTKGAGKEQDSLGQSHEVY RVKJGQVSCDIQVISDRIHICSVNESLGAAVQ LPTITIOVGNFNQTIATLQLGGSETAIVSVICSS LLLLSVVALFVFCTKSRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRIFFFRCSSLYEERYULSQTI LNSQGSSQAQETHFLLGEWKIPESCRPMMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLITLALHGKLEPKYLDSQTIL ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLCAIKQQNKGSIDAITG KARYTLNEEWLLRENIBAKRNILNVSTQGGG MOSLSVRAMDTIDTLTQVKEKILEAFCKNVFY SQWPRAEDVDLEWFASTVILRDDITSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLYDTGELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDIFS KALLSREDKPPLAVK YFFOPLEEQAKRGISD PDTHHWKTNSLPLRFWNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQGKDSPTM KLLYAKEREPYRKIVQRYKQIQDMTPLSEGE MNAHLAEESRKYQNEFNITNVAMAEIYKYAK RYRPQIMAALEANFTARRTQLQKKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ ANGKGNRYMGSPOSPLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							IAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN
CPRTLLSPLAPVPTGGSQNILVPLANTAFFOGG AALECSFGLEEIFEAVWESVNCROQVVLH TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCFAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGYWIGGVACEPLP DRYTVSEEIVCVTGAPCHSGAVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDDFTELMRIDTSIS ACTMPEGALPAPVPCVRFERRGCVHONLTF WYMQNPVITAISPRRSPSVSGRITTVAGGRFH MVQNVSMAVHHIGTPLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPOFSTAKREKWIKH HGCPLTLVHVSTIKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGG LPTTQVGNFNQTIALLJGGSSTAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGGESTAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGGSETAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPKCSSLYEERKYLPSQT LNSQGSSQAQETHPLLGGSETAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME GISLFSSLDNKHFLLJCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLVYKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKULDTERLYFFLLCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLVYKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKULDTERLYFFLLCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLSTKGTLQKFLDLDF KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLSTKGTLQKFLDDLF KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLLSTKGTLQKFLDDLF KARYTLNEEWILLRENIEARPRAFRTQLQHKFEQVVAR RCSBERKYQNEFNTNVAMAEITYXYAK RYPQIMAALEANPTARRTQLQHKFEQVVAR KLLYAKEPEYRKIVQRYYKQQFPTDGKDDVRWC VANGKGRQNTNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5747 1 382 FLKCMKKAFRSSKLLQVGYTPDGKDDVYRWC			1	!		}	GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA
CPRTLLSPLAPVPTGGSQNILVPLANTAFFOGG AALECSFGLEEIFEAVWESVNCROQVVLH TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCFAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGYWIGGVACEPLP DRYTVSEEIVCVTGAPCHSGAVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDDFTELMRIDTSIS ACTMPEGALPAPVPCVRFERRGCVHONLTF WYMQNPVITAISPRRSPSVSGRITTVAGGRFH MVQNVSMAVHHIGTPLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPOFSTAKREKWIKH HGCPLTLVHVSTIKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGG LPTTQVGNFNQTIALLJGGSSTAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGGESTAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGGSETAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELINRSC GIFFLEYKHFVTRTFPKCSSLYEERKYLPSQT LNSQGSSQAQETHPLLGGSETAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME GISLFSSLDNKHFLLJCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLVYKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKULDTERLYFFLLCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLVYKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKULDTERLYFFLLCAIKQQINKGSIDATIG KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLSTKGTLQKFLDLDF KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLSTKGTLQKFLDDLF KARYTLNEEWILLRENIEARPRNINVSFQGGG MDSLSVRAMDTDTLTKLLSTKGTLQKFLDDLF KARYTLNEEWILLRENIEARPRAFRTQLQHKFEQVVAR RCSBERKYQNEFNTNVAMAEITYXYAK RYPQIMAALEANPTARRTQLQHKFEQVVAR KLLYAKEPEYRKIVQRYYKQQFPTDGKDDVRWC VANGKGRQNTNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5747 1 382 FLKCMKKAFRSSKLLQVGYTPDGKDDVYRWC			l				QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD
AALECSFGLEUFEAVWVNESVVRCDQVVLL TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRILSDWAGWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTT WYMQNPVTIAISPRYSGGRITTVAGGER MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPOFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLGSHEY RVXIGQVSCDIQIVSDRIHICSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAITVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQREEIRKGFAGTOTDMTDLTKELNSS GIPFLEYXHFVTRTFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLLVFHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAKKNPKLMLRATESVVEKMLTNWASICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWILLRENIEAKFRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKINTLAHVFEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KALISIREDRPPLAKVFPGFASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KALISIREDRPPLAKVFPGFASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KALISIREDRPPLAKVQRYYKQQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEITYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGG VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKKMRKAFRSSKLLQVGYTPDGKDDYRWC							
TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLGPLQPMAGTCPAPEIRALEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEBIVCVTGPGPDLSQVTVVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRIDTSI ACTMPEGALPAPPVCVRFERRGCVHGNLTI WYMQNPVITAISPRRSPSVGGRTITVAGERH MVQNVSMAVHHIGFTLCKVLNSTLTICPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HHGEPL.TLVIHVSTKGAGKEQDS.LGLQSHEY RVKIGQVSCDIQIVSDRIHICSVNESLGAAVG LPITIQVGNFNQTIATI.QLGGSSTAHIVSVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELNRSC GIPFLEYKHFVTRIFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKJFSCRFNMEE GISLFSSLDDNKHFLLJCHWKJFSCRFNMEE GISLFSSLDDNKHFLLCAIKQQINKGSIDAITG KAAKPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLCAIKQQINKGSIDAITG KAAKPTLNEEWILRENIEAKPRILNYSFQGGG MDSLSVRAMDTDTLVYHALEQQKDFAVA SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHPLVJFDELAPKSH RQSHRKKVLPEFYLTVRLISTKGTLQKFLDDLF KALISREDKPPLAVKPDFLEEQAEKGGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACISIDLQLGKDSPTN KLLYAKEPEYRKIVQRYYKQQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLKTAQLPEQVSWPMGD VANGKGRQNNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMKAFRSSKLLQVGYTPDGKDDYRWC							,
VYNCAMGSPDCSQCLGREDIGHLCMWSDGG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEBIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFRRGCVHGNLTF WYMQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGGVSCDIQLOSSHIHCSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSGGSSQAQETHLPLGGWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRFSVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKLEAFCKNNPY SQWPRAEDVDLEWFASSTGSYILROLDDTSV VEDGGKKLNTLAHVPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKKVLPBIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVPDIDK TDHIDACLSVJAQAFIDACSISDLQLGKDSPTN KLLYAKEBEYSRIVQRYYKQIQDMTPLSEGE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPGIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLFEQVSWPWGQ VANGKGNQRNMGSSPQPSLLAFERNLELQIMG LGYSLLMGKLRPVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FIKKCMRKAFRSSKLLQVGYYFPDGKDDYRWC							1
RLRGPLQPMAGTCPAPERRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRSPVSGGRITIVAGERH MVQNVSMAVHHIGREPTLCKVLNSTLTICPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFLDVLPPNQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLQSHEYY RVKIGQVSCDIQVSDRIIHCSVNESLGAAVG LPITIQVGNIFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQI LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFELLCAKQQNIKGSIDATIG KARKTLNEEWLIRENIEAKPRNLNVSFQGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTIQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTI.GRVKDLDTEKYPHLVJPTDELAEPKKSH RQSHRKVLPETYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHHWKTNSLPLRFWVNILKNPGFVDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNITNVAMAEIYKYAK RYRPQIMAALEANDTARRTQLQHKFEQOVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLJEKGFLDQHKFEQOVAL MEDNIYECYSEA 731 2081 A 5747 I 382 FIKKCMRKAFRSSKLLQVGYYFPDGKDDVRWC			1				1
LLTIRGRILGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLHSLEPTIMGPKAGGTRITI HGNDLHVGSELQVLVNDIDPCTELMRTIDTSI ACTMPEGALPAPVVVRFERRSGVHGNLTT WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTILCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDVLPNPGFSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQVSDRIIHCSVNESLGAAVOC LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSI LLLISVVALFYFCKSSRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQINKGSIDATIG KARYTLNEEWILENEAKPRPLNVSFQGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTIGSYILRDLDDTSV VEDGGKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLJFDELAEPKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGIND PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNITNVAMAEITKYAK RYRPGIMAALEANTARRTGLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLHEQEVSWPWGQ VANGKGNQRNMGSSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FIKKCMRKAFRSSKLLQVGYTPDGKDDYRWC			ĺ				7
DRYTYSEEIVCVTGPAPGPLSGVYTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRIIT HGNDLHVGSELQVLVNDTDPCTELMRTIDTSI ACTMPEGALPAPVPVCVBFERRGCVHGNLTI WYMQNPVITAISPRSPVSGGRTITVAGERFH MYQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPOPSTAKREK WIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGG LPITIQVGNFNQTIATTLQLGGSGTAIIVSIVICSS LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELINRSG GIPFLEYKHFVTRTFFFRCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPIMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIAHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFILLCAIKQQINKGSIDAITG KARYTLNEEWHLRENIEAKFRNLNVSFQGCG MDSLSVRAMDTDILTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEXYFHLVLPTDELAEPKKSH RQSHRKKVLPEFYLTRLLSTKGTLQKELDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEXYFHLVLPTDELAEPKKSH RQSHRKKVLPEFYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDK KAILSIREDKPPLAVKYFFOFLEQAEKRGISD PDTLHIWKTNSLPLRFWNNILKNPQFYFDIDKDVWPGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 PSPFIFHSHETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFSSKLLQVGYTPDGKDDYRWC			ļ			į	
GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRSPVSGGRTITVAGERPH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPPPQFSTAKREK WIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHICSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSG GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSGT LNSQGSSQAQETHPLLGEWKIPESCRPIMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFTLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIGAPRRNLNVSFQGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPF SQWPRAEDVDLEWFASSTQSYLRDLDDTSV VEDGRKKLNTLAHVKPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKLVLPEIYLTRLLSTKGTIQKFLDDLF KAILSIREDKPPLAVKYFDFLEQAEKRGISD PDTLHIWKTINSLFRFWVILLKNPGFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQG MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARTIQLQHKFEQVVAL MEDNITYECYSEA QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 QFSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD							i .
HGNDLHYGSELQVLVNDTDPCTELMRITDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTT WYMQNPVTAISPRRSPVSGGRTITVAGERH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQPSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKJGQVSCDIQIVSDRIHCSVNESSGANGG LPITIQVGNFNQTIATTLQLGGSGSTAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSG GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGELEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKED NTLGRVENLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KAILSTEEMKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYTRIVLQHKFEQVSLWPWGQ MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVSLW MEDNTYECYSEA MEDNTYECYSEA VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 PSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD							l .
ACTMPEGAL PAPVPVCVREERGCVHGNLTH WYMQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGG LPITIQVGNFNQTIATILQLGGSETAIIVSIVICSI LLLSVVALPVFCTKSRRAERYWQKTLLQMB EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRIFFFKCSSLYEERYVLPSQT LNSQGSQAQETHPILIGEWKIPSSCRFNMEE GISLFSSLLDNRHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRINNVSFQGC MDSLSVRAMDIDTILTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKALDTEKYFHFILEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPGFYDIDK TDHIDACLSVIAQAFIDACISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSSQG MNAHLAESRKYQNEFNTINVAMAETYYYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				1			!
WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HIGGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRTFFFKCSSLYEERYVLPSQT LNSQGSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASTIQSYLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSREDKPPLAVKYFFDFLEPQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSFITN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 OPSPLFHSHILETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				ļ			1
MVQNVSMAVHHIGREPTLCKVLNSTLITCPSI GALSNASAPVDFFINGRAYADEVAVAEELLD PPEAQRGSRFRLDYLPPPQFSTAKREWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSC GIPFLEYKHFVTRTFFFKCSSLYEERYVYLSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFGGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEYYLTRLLSTKGTLQKFLDDLF KARLSTREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWWNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEPEYRKIVQRYYKQQDWTPLSEGG MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QFSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							ACTMPEGALPAPVPVCVRFERRGCVHGNLTF
GALŚNASAPVDFFNGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGGPLTI-UIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGC LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKBLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAALSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEPEVRKIVQAFIDACSISDLQLGKDSPTN KLLYAKEPEVRKIVQRYYKQIQDMTPLSEGE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							WYMQNPVITAISPRRSPVSGGRTITVAGERFH
PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLQSHEY RVKJGQVSCDIQIVSDRIHICSVNESLGAAVGG LPITJQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRIFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIBAKPRNLNVSFQGG MDSLSVRAMDIDTLTQVKEKIL EAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC							MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP
PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLQSHEY RVKJGQVSCDIQIVSDRIHICSVNESLGAAVGG LPITJQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRIFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIBAKPRNLNVSFQGG MDSLSVRAMDIDTLTQVKEKIL EAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC							GALSNASAPVDFFINGRAYADEVAVAEELLD
HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKJGQVSCDIQIVSDRIHICSVNESLGAAVGG LPITIQVGNFNQTIATLQLGGSETAJIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDLTQVKEKILEAFCKNVPPY SQWPRAEDVDLEWFASSTQSYHLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEFYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QFSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				i I			
RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQIDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFFLLLCAIKQQNKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKN NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							, ,
LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSY LLLLSVVALFYPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KILYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							1
LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREIRKGFAELQITDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPINEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETYGEPFFLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSIQSYILRDLDDTSV VEDGRKKLNTLAHYKPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEK YFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KARLSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKO NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QFSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							, ,
GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC	'					[,
RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFILLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTIQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
YSCLRETVGEPFFLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSIQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFFFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				[RCSLASLLTIALHGKLEYYTSIMKELLVDLID
YSCLRETVGEPFFLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSIQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFFFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							ASAAKNPKLMLRRTESVVEKMLTNWMSICM
KARYTLNEEWLLRENIEAKPRNLNVSFQGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKVDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				[·			
VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				i			
NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
TDHIDACLSVIAQAFIDACSISDLQLĞKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC						, 1	
MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC						, i	TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN
MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE
RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							•
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	. I]	!		
VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	720	2000		5744		202	
LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD	/30	2080	А	3/44	3	292	
731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				'			
	731	2081	Α	5747	1	382	FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC
	·]]			FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

NO: of peptide colide sequence uence	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
sequence Sequence	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
uence 914 ng to first amino acid residue of peptide residue of peptide residue of peptide sequence 92 peptide sequence 92 peptide sequence 93 peptide sequence 94 peptide sequence 94 peptide sequence 94 peptide sequence 94 peptide sequence 94 peptide sequence 95 peptide sequence 94 peptide sequence 95 peptide 95 peptide sequence 95 peptide seque	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
amino acid residue of peptide sequence peptide sequence peptide sequence se	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
residue of peptide sequence	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
peptide sequence Possible nucleotide deletion, =possible nucleotide insertion	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
Sequence Incleotide insertion LSFSLRSSRVSGRHWKNFALVPLLREAS. RQSAQPEEVYLRQFSGSLKPEDAEVFKS GEK 732 2082 A 5753 198 3 AQAESSTVASPEATAGPLCTRIPNVPPPT 733 2083 A 5754 2 2223 AAGPPGLAEGRAPESAGPGPGGDAAET PPAHSGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEEEETAEVGRGAEAGAGA LNRYTSTSTKSAKSGSEASASASKDALQA LPRYHCHAPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFINRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVE AFSQRYCMCNPEVVQQFHNPDTIFILAF NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALREPAAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGGGLEGQIVARDENSVLIDG DDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
LSFSLRSSRVSGRHWKNFALVPLLREAS. RQSAQPEEVYLRQFSGSLKPEDAEVFKS GEK 732 2082 A 5753 198 3 AQAESSTVASPEATAGPLCTRIPNVPPPT PGKLQAQLPCPSPVRFTSARIPPASRPQTT 733 2083 A 5754 2 2223 AAGPPGLEAEGRAPESAGPGPGGDAAET PPAHSGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGA. ATAPKTEEEFEEEFTAEVGRGAEAEAGL LNRTSTSKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF, NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHIRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKGSPTAKGAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPTTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
RQSAQPEEVYLRQFSGSLKPEDAEVFKS. GEK 732 2082 A 5753 198 3 AQAESSTVASPEATAGPLCTRIPNVPPPT PGKLQAQLPCPSPVRFTSARIPPASRPQTI 733 2083 A 5754 2 2223 AAGPPGLEAEGRAPESAGPGPGGDAAET PPAHSGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGA, ATAPKTEEEEEEEETAEVGRGAEAEAGL LNRTSTSTKSAKSGSEASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKPQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF, NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCFK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPFTDTLVQCQQIVKVIVLL LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	AAS PIRP S TAL GGR DG LEQ HILS GLN QRK EM RLIE HILL DG
732 2082 A 5753 198 3 AQAESSTVASPEATAGPLCTRIPNVPPPT 733 2083 A 5754 2 2223 AAGPGLEAGRAPESAGPGPGGDAAET PPAHSGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTPQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEEEETAEVGRGAEAEAGL LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKPQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPTPFTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPTYNHEHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	PIRP S S PGL TTAL S G R D G L E Q G I I I S G R L E G I I I I I I I I I I I I I I I I I I
732 2082 A 5753 198 3 AQAESSTVASPEATAGPLCTRIPNVPPPT PGKLQAQLPCPSPVRFTSARIPPASRPQTT 733 2083 A 5754 2 2223 AAGPPGLEAEGRAPESAGPGPGGDAAET PPAHISGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGA. ATAPKTEEEEEEETAEVGRGAEAEAGLLNRTSTSTKSAKSGSEASASASAKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRJ LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI) DFSSMELDEALRKFQAHIRVQGEAQKVV AFSQRYCMCNPEVVQQFHNPDTIFILAF, NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPPPPPTPTGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG DDVMRVLDDIDFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	PGL TTAL GGR DG LEQ MILS GLN QRK EM RLIE IILL DG
PÖKLQAQLPCPSPVRFTSARIPPASRPQTT 733 2083 A 5754 2 2223 AAGPPGLAEGRAPESAGPGPGGDAAET PPAHSGTILMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEEEETAEVGRGAEAEAGL LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG UGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHIFWSLGAGRGAPVPPPDLQF QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPTPNHPHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG DDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	PGL TTAL GGR DG LEQ MILS GLN QRK EM RLIE IILL DG
733 2083 A 5754 2 2223 AAGPPGLEAEGRAPESAGPGPGDAAET PPAHSGTILMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEEEETAEVGRGAEAEAGL LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVF AFSQRYCMCNPEVVQQFHNPDTIFILAFANTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHOT TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPNHPHQFCPPGS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVQQGLEEQIVARDENSWLIDC IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	PGL TAL GGR DG LEQ MILS GLN QRK EM RLIE IILL DG
PPAHSGTLMMAFRDVTVQIANQNISVSS SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPA VGREDASAEDSCAEAGASGAA ATAPKTEEEEEEETAEVGRGAEAEAGE LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVE AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALREPAESTV HNRLQTSQHNSGLGAERGAPVPPDLQP QQTPPLPPPPPPPFPTFGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFFQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	TAL GGR DG LEQ MILS GLN QRK EM RLIE IILL DG
SVANCLGAQTVQAPAEPAAGKAEQGET EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEETAEVGRGAEAEAGL LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF, NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHINSGLGAERGAPVPPPDLQF QQTPPLPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	GGR DG LEQ MILS GLN QRK EM RLIE HILL DG
EAPEAPAVGREDASAEDSCAEAGASGAA ATAPKTEEEEEEETAEVGRGAEAEAGE LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVF AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKX QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFFQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	DG LEQ MILS GLN QRK EM RLIE IILL DG
ATAPKTEEEEEETAEVGRGAEAEAGIL LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	LEQ MILS GLN QRK EM RLIE IILL DG
LNRTSTSTKSAKSGSEASASASKDALQA LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAF, NITDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFK QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	AILS GLN QRK EM RLIE IILL DG
LPRYHCENPASCKSPTLSTDTLRKRLYRI LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVI AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPFGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	GLN QRK EM RLIE IILL DG
LFNINPDKGIQFLISRGFIPDTPIGVAHFLI GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVE AFSQRYCMCNPEVVQQFHNPDTIFILAF NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHO TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	QRK EM RLIE IILL DG
GLSRQMIGEFLGNSKKQFNRDVLDCVVI DFSSMELDEALRKFQAHIRVQGEAQKVE AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPTPPGTLVQCQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	EM RLIE IILL DG
DFSSMELDEALRKFQAHIRVQGEAQKVE AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGJYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPTPPGTLVQCQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	RLIE IILL DG
AFSQRYCMCNPEVVQQFHNPDTIFILAFA NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKJ QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	IILL DG
NTDMYSPNIKPDRKMMLEDFIRNLRGVI ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	DG
ADIPRELVVGIYERIQQKELKSNEDHVTY VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
VEKSIVGMKTVLSVPHRRLVCCSRLFEV NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEL SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
NKLQKQAAHQREVFLFNDLLVILKLCPK SSTYTFCKSVGLLGMQFQLFENEYYSHG TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPTGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
SSTYTFCKSVGLLGMQFQLFENEYYSHO TPLSGSEKKQVLHFCALGSDEMQKFVER SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPPFPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
TPLSGSEKKQVLHFCALGSDEMQKFVEI SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
SIAEVTELEQIRIEWELEKQQGTKTLSFKI QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPTPPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
QGDPQSKQGSPTAKREAALRERPAESTV HNRLQTSQHNSGLGAERGAPVPPPDLQP QQTPPLPPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSI GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
HNRLQTSQHNSGLGAERGAPVPPPDLQF QQTPPLPPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSS GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
QQTPPLPPPPTPGTLVQCQQIVKVIVLI LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
LARMEPLLSQALSCYTSSSSDSCGSTPLG SPVKVTHQPPLPPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
SPVKVTHQPPLPPPPPYNHPHQFCPPGS: GHRYSSGSRSLV 734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
734 2084 A 5788 8 362 SSVMGDLVGQGLEEQIVARDENSWLIDG IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	
IDDVMRVLDIDEFPQSGNYETIGGFMMF KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	GTP
KIPKRTDSVKFAGYKFEVVDIDNYRIDQI	/ILR
735 2085 A 5827 I 1257 MVFSAVLTAFHTGTSNTTFVVYENTYM	ITL,
PPPFQHPDLSPLLRYSFETMAPTGLSSLT	NST
AVPTTPAAFKSLNLPLQITLSAIMIFILFV	FLG
	DM 1
LLAVLNMPFALVTILTTRWIFGKFFCRVS	AMF
AKVLIAVSWATSFCVAFPLAVGNPDLQI	
PQCVFGYTTNPGYQAYVILISLISFFIPFLV	
SFMGILNTLRHNALRIHSYPEGICLSQASI	
MGLQRPFQMSIDMGFKTRAFTTILILFAV	
WAPFTTYSLVATFSKHFYYQHNFFEISTV	
WLCYLKSALNPLIYYWRIKKFHDACLDN	
KSFKFLPQLPGHTKRRIRPSAVYVCGEHI	
736 2086 A 5870 3 268 FTRSDELARHYRTHTGEKRFSCPLCPKQI	
DHLTKHARRHPTYHPDMIEYRGRRRTPR	DPP
LTSEVESSASGSGPGPAPSFTTCL	
737 2087 A 5871 2 521 LTWPQLFLETLPELLHMSRPAEDGPSPGA	
RSSSLGYISKAEEYFLLKSRSDLMFEKQS	RH !
GLARRLTTARRPPASSEQAQQELFNELKI	
DGANFIVNHMRDQNNYNEEKDSWNRV	ΑV
VDRLCLFVVTPVMVVGTAWIFLQGVYN	AV RT
QPFPGDPYSYNVQDKRFI	AV RT
738 2088 A 5881 I 1160 LVVTAITAILAFPNEYTRMSTSELISELFN	AV RT)PPP
LLDSSKLCDYENRFNTSKGGELPDRPAG	AV RT PPP
YSAMWQLALTLILKIVITIFTFGMKIPSGI	AV RT ()PPP OCG 'GV
MAVGAIAGRLLGVGMEQLAYYHQEWT	AV RT PPP CG GV FIPS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WCSQGADCITPGLYAMVGAAACLGGVTRMT VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL AMDVMKPRRNDPLLTVLTQDSMTVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI LFN
739	2089	A	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP DQALQELRKVARINGHKEAKNLTIEVLMSSV KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA VDFLGRATTALLLSFLGRRTIQAGSQAMAGL AILANMLVPQDLQTLRVVFAVLGKGCFGISL TCLTIYKAELFPTPVRMTADGILHTVGRLGA MMGPLILMSRQALPLLPPLLYGVISIASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM
741	2091	А	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV
742	2092	A	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC
743	2093	A	5938		1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRLL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
745	2095	A	5970	413	856	LYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL NFPRRAQGPTWRSLEVIQSTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	""		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
defice	i		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Ì	[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	•	1	ľ	peptide	sequence	/=possible nucleotide deletion, \=possible
1	1	}			}	nucleotide insertion
246	2096	 	6071	sequence 3	1040	
746	2096	Α	5971	۱ ،	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
	ļ	ļ				RCARHGACQRSCLASQDPYCGWHSSRGCVDI
1	1	1			1	RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
		}				SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
1	ĺ	ľ		i	i	AAAFALGASVSGLLVSCACRRAHRRRGKDIE
1						TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
i		1	ĺ		İ	VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
						LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
	ţ	ŀ		[GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
Į]	l			ĺ	EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
		1			ļ	APALLGGPSPRPHECASPLRLDVPPEGRCASA
ĺ		ĺ		1	İ	PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
ŀ		l				LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
		1				RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG
}	ļ	l	1			
747	0007	<u> </u>	5000		554	GRFNF
/4/	2097	Α	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
i						LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
1		ţ.				RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
Ì		j				DRVQALSYAQHTRQLISCGGDGGIVVWNMD
						VERQETPEWLDSDSCQKCDQPFFWNFKQMW
						DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
1						MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
1 .						HNIVHVHFDATRGWLLTSGTDKVIKLWDMT
						PVVS
748	2098	Α	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
•				-		CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
i .						TMLLMLKLCTEVRVANELNARRRSFTAADS
1						KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
			}			QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
1			1			LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
			!			WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
			1 1			DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
749	2099	Α	6002	2	447	
149	2099	A	0002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
						RARLADDLNEKIAQRPGPMELVEKNILPVDSS
						VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
		ľ				DQPASQESQGSAASPSEPKVSESPSPVTTNTP
						AQFASVSPTVPEFLKTPPTAD
750	2100	Α	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
)]			WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
1				ļ		QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
						WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
]			GQEDSLDSAVDAATEHKTC
751	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
	_	-		_		SHPDKLKRMSKSVPAFLQDESDDRETDTASE
						SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
						VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
						HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
1			}			DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
1						QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
j l				_		TWDWDNKQNKQLRWYPLKRKTAPVALEAE
						NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
				1		KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
						KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
				i		VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
<u> </u>				l		EVDWMDSTSEEVALWEKMVNSPNTWIEATL
L						PLRMLLIAKISK
752	2102	Α	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL
				į		SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF
}				į	J	AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDGSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK
753	2103	A	6043		1470	KNLKEKA DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	Α	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRFTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE- GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEGEEPNYPLELLYCLYETQEDAFVRQA LCRFFELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106	A	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	Α	6063	54	419	ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į		1		peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
750	0100	<u> </u>	6076	100		QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	Α	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
		1		ł		PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
						LRAVLKLMSECWAHNPASRLTALRIKKTLAK MVESQDVKI
759	2109	A	6072	3	650	PGRRFRPAALEERAMEKLREKVPFONRGKGT
139	2109	^	0072	,	030	LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
1 .	İ					SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA
		İ				LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP
						KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV
						HSSVADMONMPAAVHALLTOPSLSAAPFAO
	[RYLGTLPSTGSTTLPQCHAGNATVW
760	2110	Α	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG
	ļ					LRGCMLIELPLRGRLQLEACGMRRKSLLTRK
1						VICKSDAPTGDVLLDEALKHVKETQPPETVQ
]						NWIELLSGETWNPLKLHYQLRNVRERLAKNL
1						VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR
	i					LIKKVQEAVLDKWVNDPHRMDRRLLALIYL
						AHASDVLENAFAPLLDEQYDLATKRVRQLLD
L						LDPEVECLKANTNEVLWAVVAAFTK
761	2111	Α	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
1						HQKLSADMADHSNLIRSLLVGAEDARLMRD
1	ĺ					MKTMKSRYMELYDLNRDLLNGYKIRWNNH
						TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
762	2112	Α	6079	2	2686	ACRDAIRSNNINTLFKIMRVGTASS
/02	2112	А	0079	2	2080	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG
	i 1					SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI
į						QLAKLDSEAWPGVLDSERDRLILINEKEELLK
}						EMRFISPRK WTQGEVEQLEMARKRLEKDLQ
1					:	AARDTOSKALTERLKLNSKRNOLVRELEEAT
[RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR
						GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ
						SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ
ĺ						KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS
						PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA
						TLCELSLGNSAQERYRLEEPGTEGKQLGQAV
						NTAQGCGLKVACVSAAVSDESVAGDSGVYE
						ASVQRLGASEAAAFDSDESEAVGATRIQIALK
						YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR
						VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW
1						VSMSYPALHQKTLRVDVCTTDRSHLEECLGG
			ĺ			AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEOTAVE
						LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
						EEEEEVEEEGEEDVFTEKASPDMDGYPALK
						VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF
} l						LRGSTHRSKTFSPGPQSQYVCRLNRSDSDSST
						LSKKPPFVRNSLERRSVRMKRPSPPPOPSSVK
 						SLRSERLIRTSLDLELDLOATRTWHSOLTOEIS
		J			` \	VLKELKEQLEQAKSHGEKELPQWLREDERFR
j		ļ	-	į		LLLRMLEKRMDRAEHMGELQTDKMMRAAA
		ł	ŀ			KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR
1						MNIPALSADDV
			I			
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}	l	ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
			Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ľ	!		1	peptide	,	/=possible nucleotide deletion, \=possible
}	i		l	sequence		nucleotide insertion
—				55400.00		VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK
						THVTLHGTELCDESYPALLTDIPVGDLHPGEO
İ	•					LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE
Į.]		1			KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF
						EHLERVYADIPFLLMTDLLSASPWALTIVSSE
İ	į	ļ				LHLAPSMTTVDQLESQVDNVILQTGESASECF
	ļ					CLOCPSLGNIEGGVATGHYIISWKRTSAMENI
1	1	1				PIITTVITLPHVIVENIPLHVNADLPSFGRVRES
}	i I					
İ						LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG
ŀ						LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
		1				LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
7/4	2114		(002		1400	DDTSIAAA
764	2114	Α	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP
						PPPAPAPPTLGNNHQESPGWRCCRPTLRERN
						ALMFNNELMADVHFVVGPPGATRTVPAHKY
1						VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA
i						AFLILLKYMYSDEIDLEADTVLATLYAAKKYI
1						VPALAKACVNFLETSLEAKNACVLLSQSRLF
1	!					EEPELTQRCWEVIDAQAEMALRSEGFCEIDR
						QTLEIIVTREALNTKEAVVFEAVLNWAEAEC
ĺ		ĺ				KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE
1						FANGAAQSDILTLEETHSIFLWYTATNKPRLD
į						FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG
ł						RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV
j						KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF
		:				EHPVQVEQDTFYTASAVLDGSELSYFGQEGM
						TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE
				,		LIFYA
765	2115	A	6099	1	1150	SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF
1		·				RPVKAPGTFHMVHGKCMCKHNTAGSHCQH
					•	CAPLYNDRPWEAADGKTGAPNECRTCKCNG
						HADTCHFDVNVWEASGNRSGGVCDDCQIIN
J						TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS
						CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA
						GRRCDRCMVGYWGFGDYGCRPCDCAGSCD
						PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP
						WEWEDAQGFSALLHSGKCECKEQTLGNAKA
1		l				FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK
					!	KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL
						NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK
						PSLGRKVMDILKRECK
766	2116	Α	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR
<u>†</u>					İ	CVLTERGLQLFEAKGTGGRPKELSFARIKAVE
						CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW
1						NAQITLGLVKFKNQQAIQTVRARQSLGTGTL
						VS
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP
		·		-	- · -	MARYYIIKYADQKALYTRDGQLLVGDPVAD
						NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC
						LACVETEEGPSLQLEDVNIEELYKGGEEATRF
(TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ
]						PVQLTKESEPSARTKFYFEQSW
768	2118		6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF
/06	4110	Α	0103	3	272	
J						YTPSVISSVMHRVARCAAPHVHILLANFYLLF
	2116					PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A	6110	. 1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSS
[SSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ
			1			ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY
L						KHEDLQTDESSMDDRHPRRQLCGGNQAATE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RIILFGRELQALSEQLGREYGKNLAHTEMLQD AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
770	2120	A	6125	2	570	NSAILESQNLPKQPPLMLALGQASECLRLMA RAGLGSCSFARVDDYLH YFGLNLHVQHLGNNVFLLQTLFGAVILLANC VAPWALKYMNRRASQMLLMFLLAICLLAIIF VPQEMQMLREVLATLGLGASALANTLAFAH
771	2121	A	6126	909	353	GNEVIPTIIRARAMGINATFANIAGALAPLMM ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK PLFDTIQDEKNERKDPREFKQEDPRVEVTQF RSFVLDTASAICNYNAHYKNHPKYWCRGYF
		A	0123	909		RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC LTKEDTGWYWCGIQRDFARDDMDFTELIVT DDKGTLANDFWSGKDLSGNKTRSCKAPKVV RKADRSRTSILIICILITGLGIISVISHLTKRRS QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	А	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV WYLLRKHWIANNLFGLAFSLNGVELLHLNN VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV ALAKGEVTEMFSYEESNPKDPAAVTESKEGT EASASKGLEKKEK
773	2123	A	6161	3	1088	CQPMLVTRKNHPKLLLRRTESVAEKMLTNW FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV NPENENAPEVPVKGLDCDTGTQAKEKLLDA AYKGVPYSQRPKAADMDLEWRQGRMARIIL QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH LDQREGDRGSKMVSEIYLTRLLATKGTLQKF VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ ADKHQIHDADVRHTWKSNCLPLRFWVNVIK NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA GSEGAGLPPSGELHFWVKEARDLLPLRAGSL DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF NHTMVYDGFGPADLRQACAELSLWDHGALA NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVQ \PWNWLMLGCHTAVDFDQLISSMPCISHGMT ASASAL
776	2126	A	6217		827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT RQKHAKKHLGFFRNNFGVREPYQILLDGTFC QAALRGRIQLREQLPRYLMGETQLCTTRCVL KELETLGKDLYGAKLIAQKCQVRNCPHFKNA VSGSECLLSMVEEGNPHHYFVATQDQNLSVK VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA VESGRLSQCMRKKVSNISKRNRV**KTLNRG RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE KKRKRKRIRNRSNPKVLSEKQNAEGE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL* FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	Α	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN AVSNAD\LLDLK\N*LDH\LEEKMPLIEVKVVP PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/ SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC NSFRYRR
779	2129	Α	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP YGQSQPSCFDRVKMGFVMGCAVGMAAGAL FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH QSQPMY
780	2130	A .	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG TDYWLYSRGVCRTKSTSDNETSRKNEEVMT HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV
781	2131	A	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN
782	2132	Α	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKTKKKKNKKTKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	Α	6305	201	1032	WDDYPQGALRRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSI.I.SAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAALSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFRFEFWDCGESA LKKFDHMI.LACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPIPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RSKLESLCRELQRHNRSLKEEGVQRAREEEE KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR QENMELAERLKKLIEQYELREEHIDKVFKHK DLQQLVDAKLQQAQEMLKEAEERHQREKD FLLKEAVESQRMCELMKQQETHLKQQLALY TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI
705	0125		(210	1402	200	KKLEKETTMYRSRWESSNKALLEMAEEKTV RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR GQRWGSHRTSAVRIFS
785	2135	A	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT KLPWSWGMRPMKIFFSEEYRSISTRISHDAL* EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVRKNENCYNTP GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSL*NSLJ*AVTSYEDLGLFAFGLPGKL VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMT\V KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	Α	6351		6622	PRSLCFSLWAEAAVLADGGLRRRRILLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMTYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNILKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

	Cane as			-		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ŀ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
						RQNSEREAGKKHKVEDGTSSVTVLSHEEDA
		i				MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
J		l	1		J	VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
		l	1			ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
			ļ			EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
İ		1				GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
i		ŀ				l
i						KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
		ļ	1			CTYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
			i			DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
		ļ	j		Į.	LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
!						VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
						EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
į						NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
						MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
						KRFEALRATSAEHVNGGLKESEKLSDDLILLL
					i	QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
						EERDCCNDCYLALEHGRQFMDNMSGGKVDE
.]]			ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
						LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
						KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
			F			LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
ı						· · · · · · · · · · · · · · · · · · ·
1				•		NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
			}			LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
						GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
J)	1			SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
l			1			ATLASAASSTMTVTSGVAISTSVATANSTLTT
i			1	į		ASTSSSSSNLNSGVSSNKLPSFPPFGSMNSNA
į			1			AGSMSTQANTVQSGQLGGQQTSALQTAGISG
Ì						ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH
!						AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
į						GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
i						PVKHEDREIYPOHLKSLAFSAFTOCRRPLPTS
						TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
]			PPFILAPVKDKQTELGETFGEAGQKYNVLFV
			1			GYCLSHDORWILASCTDLYGELLETCIINIDVP
ļ						NRARRKKSSARKFGLQKLWEWCLGLVQMSS
			1			LPWRVVIGRLGRIGHGELKDWSCLLSRRNLO
			1			•
ŀ						SLSKRLKDMCRMCGISAADSPSILSACLVAM
						EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
į			J !			NTPQDTSCTHILVFPTSASVQVASATYTTENL
						DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
ļ						NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
			[RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
•						LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
ļ				j	İ	QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
ļ			l			ALSWLTCDPATODRRSCLPIHFVVLNQLYNFI
i		l	[]			MNML
789	2139	Α	6359	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
			5557	^	2002	LPVGPLLRALATCHALSRLQDTPVGDPMDLK
Ì						MVESTGWVLEEEPAADSAFGTOVLAVMRPP
İ						
1						LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
1						SVVVAWPGATQPEAYVKGSPELVAGLCNPET
İ						VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
		l				SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
						QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
					l	RGCGMVAPQEHLIIVHATHPERGQPASLEFLP
I						MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
			1			MINDI IN THE TRUE DESCRIPTION OF THE PERSON
						-
				-		LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
				-		-

CCC 75	Locom	14-4	Lega	Dendicted	D. 4: 4 - 1 - 1	I Amine poid cogness (Am Alexino C-Courte)
SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of peptide	hod	in NO:	nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide			USSN	location		l=Isoleucine, K=Lysine, L=Leucine,
	seq- uence		09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	,	Y=Tyrosine, X=Unknown, *=Stop codon,
					sequence	/=possible nucleotide deletion, \=possible
	ļ			peptide sequence	1	nucleotide insertion
		<u> </u>	 	sequence		VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL
j .	1			ł		
						YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP
1	İ				ļ	ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG
į.		i			i	VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY
		İ		ł	ļ	ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN
						NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR
l	i	İ	İ			NITDTGFKLLLVGLVTLNFVGGLHAGERARP
1	1	· .			<u> </u>	VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW
		L				PPLPAGPLR
790	2140	Α	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP
						GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT
		ļ			İ	FKRGLLLSALSYLGFETYQVISQAAVVHATA
		1	1		l	KVEEILEQADYLYESGETEKLYQLLTQYKESE
	1	l	1	1		DAELLWRLARASRDVAQLSRTSEEEKKLLVY
1	1	l		1		EALEYAKRA/L/EKNESSFASHKWYAICLSDV
f						GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS
					į	IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP
						*FPPYEKALG\YFHRAEQVDPNFYSKNLLLLG
1		1			i	KTYLKLHNKKLAAFWLMKAKDYPAHTEED
1		ŀ				KQIQTEAAQLLTSFSEKN
791	2141	Α	6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS
	l					SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ
1					ļ	*VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF
i	I	i				AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ
İ						RLQEQRQQQSGEAEALARVYSSSISNGLSNLN
		, i				NETSGTYANGSVIDLPKSEGYYNVVSGQPSP
	l					DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS
		ł				FNN\GQLAPGIT\MTEIDRIAQNIIKSHLETCQY
İ	ĺ					TMEELHQLAWQTHTYEEIKAYQSKSREALW
	1					QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ
ļ ·						ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG
	ŀ					KYGGMQMFKALGSDDLVNEAFDFAKNLCSL
1	į.					QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ
						EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA
1	ŀ		!			VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF
	L		L			NPDCATACK
792	2142	Α	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV
1		l				TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL
1	1	1				MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT
	l	-				LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA
]]]			LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH
[1					ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH
	1					GAVVNESHHDALVEDIFDKEDEDKDGFISAR
	l ·					EFTYKHDEL
793	2143	A	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPAELG
	I					ARADQELVTALMCDLRRPAAGGMMDLAYV
						CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM
			1			DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH
1						EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS
			[MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\
						WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS
						P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLO\P
						SGOVL\TST\ESLCRLRARVALADIAFTGGGNI
						VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT
						DILPSLFMRCTTDLNRKDKFPAITHLKFLARD
1	1					I .
	1					MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI
	1	}				FQQISPVVGDKQPTILKWRILSATNDLDRVSA
1	1]			VALPKLPISLTNTDLKVASDTQFYPGLGLAL
}	1		[AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD
L	L	L	L	L		EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IDSHGKLSVLRLSPSMGHPLEVGLALRHLLFL LEYCMVTGYDWWDILLHVQPSMVQSLVEKL HEEYTRQTAALQQVLSTRILAMKASLCKLSP CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLCG SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC RDEGPASEPDEALVDECCLLPSQLLIPSLDWL PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT LQLDGLARAPGQPKIDHLRRLHLGACPTEEC KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
794	2144	A	6490	418	585	NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT PRSLDHLHPEDRP NGDKADLENESCRAQVLMPVVPALWEAEGG
						GSIEPRDLRLQ*AVITPL\TPAWVTQ
795	2145	Α	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS DTRVRAAVGAPGLPVEPLV
796	2146	Ā	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM KKMNKLQNISEELQRNISLQLMSNMNISNKIR NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR WIWHKDSCYFLSDDVQTWQESKMACAAQN ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE DSYYSWYESG*YNQVPSAWVIRNAPDLNNMY CGYINRLYVQYYHCTYKQRMICEKMANPVQ LGSTYFREA
797	2147	A .	6507		881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH APRAASAQLEERMRDPHPGMTLQEGDCRGS QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA AIWTGAVAVLAGAAAFIYEKRGGTYWALLR TLLALAAFSTAIAALKLWNEDFRYGYSYYNS ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM DMLKALFRTLQAMLLGVWILLLLASLTPLWL /SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG STHASARSQVPRSAGEAAPHSRRPPGLLPHAP RAASAQLEERMRDPHPGMTLQEGDCRGSQT VSLTMGTADSDEMAPEAPQHTHIDVHIHQES ALAKLLLTCCSALRPRATQARGSSRLLVASW VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI WTGAVAVLAGAAAFIYEKRGGTYWALLRTL LALAAFSTAIAALKLWNEDFRYGYSYYNSAC RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM LKALFRTLQAMLLGVWILLLASLTPLWLYC WRMFPTKGVSP
798	2148	A	6528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP RFLVAFAYWNHYLSCTSPCSCYRPLCKLNFG LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV MVRWELSSNGNPGRGVLGLGLGLGNKLRVV GQNLGL*HCVWVVWETGE*KRWRLQMGIE*GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL GPSLPQRQGREHIVVILAAPACAPFHDR*WEP REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D RKSYSWKQRLFINFISFFSALAVYFRHNMYC
799	2149	A	6529	1	874	EAGVYTIFAILEYTVVLTNMAFHMTAWWDF GNKELLITSQPEEKRF FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
						CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIIEGSEREHEETRTKQAALDGEPLGGGQLTA VHLHPSKEQQGQEGGERQRGARTHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK
800	2150	A	6544	2		SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG E\THLCFLVR\DRVSALTQMESACVSIHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
801	2151		6556	1	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR
802	2152	A	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL FAKLKTQTNASDFPIGTSLKYECRPEYYGRPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS TNRENFHYGSVVTYRCNPGSGGRKVFELVGE PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP RRVKCQALNKWEPELPSCSRVCQPPPDVLHA ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
						QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
						GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
		·				KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP
803	2153	Α	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DGLLLYIQAVTETLAHGGTVTDGENITQRMW
						NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE NGAFRVVLNYNGTSQELVAVSGRKLNWPLG YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV GSLSLLGILIVSFFIYRKMQLEKELASELWRVR WEDVEPSSLERHIRSAGSRLTLSGRGSNYGSL LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR IELTRKVLFELKHMRDVQNEHLTRFVGACTD PPNICILTEYCPRGSLQDILENESITLDWMFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV DGRFVLKITDYGLESFRDLDPEQGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGIILQE IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFQ QIRLTLRKFNRENSSNILDNLLSRMEQYANNL EELVEERTQAYLEEKRKAEALLYQILPHSVAE QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE STPMQVVTLLNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRLHACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSYETKAVLYEEFGGFELELRGDVEMKGKG
804	2154	A	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEGEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	ucileo		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						NRLWVGTGNGVIISIPLTETVILHQGRLLGLR ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
						FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
						ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	Α	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
	,					SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
						SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLO
						DSGLYACVIRNSTYCMKVSISLTVGENDTGL
						CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
						REPEILWYKECRTKTWRPSIVFKRDTLLIREV
						REDDIGNYTCELKYGGFVVRRTTELTVTAPL
			}			TDKPPKLLYPMESKLTIQETQLGDSANLTCRA FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
						SDI\KILKEHLGEQEVSISLIVDSVEEGDLGNYS
}						CYVENGNGRRHASVLLHKRELMYTVELAGG
						LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA
						EELDGDNKDYDAYLSYTKVDPDQWNQETGE
						EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF
1				•		ELETRLRNMLVTGEIKVILIECSELRGIMNYOE
						VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
						LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
			•			VSAISMAAATSTALATAHPDLRSTFHNTYHS
						QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
:						ILPLLPRETSISSVIW
806	2156	A	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL
						HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT
}						QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL
	•					IQNKYFGDVDIPRAKVVRVCQALMDYKVFE AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNOD
						SQLGKENKLYSPARYADALFKSSDIRSASLED
						LWENLSLKPANSPHVNISTTLSPQVINEVWQE
						ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
	i					RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA
						AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF
						MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
						DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT
1						L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
						KEKKK\LLGQFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
						PQILLLLALLTLGLAAQHQDKVPCKM/VKML
						CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD
						LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG
						GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS
						LHTLSLAENSLTRI.TRHTFRDMPALEQLDI.HS
						NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD
						FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT
						WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR
						LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL
			}			NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE
						TLELGARALG\SLRTLLLQGNALRDLPPYTFA
					ĺ	NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\
Li]	AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LDLSSNPGLEVATGALGGLEASLEVLALQGN GLMVLQVDLPCFICLKRLNLAENRLSHLPAW TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA
808	2158	A	6619	153	1852	TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI NLIILTFILVSAILLTTLAACCCVRRQKFNQQ YKA FKALSQYIYTNTHLEREAAFEVAILLRRMEEG
·						ARHRNNTEKKHPGGGESDASPEAGSGGGGV ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN AGSVGLALIVWIVTGFITVVGALCYAELGVNI PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLLTWVNCSSVRWATRVQDIFTAGKLL ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP DIGLVALAFLQGSFAYGGWNFLNYVVTEELV DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP SVLAMIHVKRCTPIPALLFTCISTILLMLVTSD MYTLINYVGFINYLFYGVTVAGQIVLRWKKP DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI ELLTLVSQKMCVVVYPEVERGSGTEEANED MEEQQQPMYQPTPTKDKDVAGLOVAEGVPI
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV VLIDCVGVGLLIATLMWFISNKYLVKRQSRD YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL GYSVGLLFFSVALPFLKNTVILLYPFAPLILLYG LSLALGWNFTHTLCSFYKYRVK
810	2160	Α	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT QKMLDNFYNFASSFAVSQVPDDTQ/RPSEMF IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	A	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG AAASISQEPSLADIKARIQDSDEPDLSQNSITG EHSQLLDDGHKKARNAYLNNSNYEEGDEYF DKNLALFEEEMDTRPKVSSLLNRMANYTNLT QGAKEHEEAENITEGKKKPTKTPQMGTFMG VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF AIVLICCCCTMLTAISMSAIATNGVVPAGGSY FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGHSKANGEPTWALLLTAAIAELGILIASLDL VAPILSMFFLMCYLFVNLACALQTLLRTPNW
						RPRFRYYHWALSFMGMSICLALMFISSWYYA IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ LVVAAKLREGISHLIQSCGLGGMKHNTVVM GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG GMLMLLPFLLK\QHKVWRKCSIRFFTVAQLE DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS DISAYTYERTLMMEQRSQMLRHMRLSKTER DREAQLVKDRNSMLRLTSIGSDEDEETETYQ EKVHMTWTKDKYMASRGQKAKSMEGFQDL LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL ERVLLVRGGGSEVITIYS
812	2162	Α	6628		640	AVCTMSEMAELSELYEESSDLQMDVMPGEG DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG CDEIIDRE
813	2163	Α	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC WQYRQLSALHRAPRPTRPDKARRLGYKAKQ GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG ATYGKPVHHGVNQLKFARSLQSVAEERAGR HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS RGLGKGHKFHHTIGGSRRAAWRRNTLQLH RYR
814	2164	A		201	1705	KGTEMNKSR WQSRRRHGRRSHQQNPWFRLR DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN PLTKESIRQKEMESKRLRLLQEEDRRKKIARM GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF MHENLYFTNRKVNSVCWASLNHLDSHILLC LMGLAETPGCATLLPASLFVNSHPAGIDRPG\ MLCSFRIPGA WSCAWSLNIQANNCFSTGLSR RVLTNVVTGHRQSFGTNSDVLAQQFALMA PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

000 00	1 000 10	136.	Leno		I 8 - 17 - 1	-
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *≈Stop codon,
	Ì			peptide		/=possible nucleotide deletion, \=possible
		ļ	<u> </u>	sequence	ļ	nucleotide insertion
i				ł		RVFETLKDLKVLNLAYNKINKIADEAFYGLD
		1				NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
1	}	Ī		Į.		QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
				ļ		FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR
ļ]]	}	j	ļ	LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ
		l			ł	TPSENPSLEQLFLGENMLQLAWETELCWDVF
			Ì			EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR
						GLSLNSNRLTVLSHNDLPANLEILDISRNQLL
						APNPDVFVSLSVLDITHNKFICECELSTFINWL
i		ł	ľ		ľ	NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE
			ļ	ĺ		GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV
						TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY
-						KYDAYLCFSSKDFTWVQNALLKHLDTQYSD
Ī						QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
		}				KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL
		Ì	-			NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ
!				ŀ		YLRWPEDLQDVGWFLHKLSQQILKKEKEKK
016	01//	ļ. <u>.</u>	1	ļ		KDNNIPLQTVATIS
816	2166	A	6646	1	3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
Ì	1					GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
			1			GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
	İ		l			KKGQGTAATGNQATPKTAPATMSTPTILVAT
ł			İ			AVHAYRYTNGQYVKQGKFGAAVLGNHTTR
			١,			EYRILLYISQQPVTVARIHVNFELMVRPNNY
Ì						STFYDDQRQNWSIMFESEKAAVEFNKQVCIA
						KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE
ļ]				VAYTGWLFQNHVLGQVFDSTANKDKLLRLK
			}			LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA
		Ì	İ			KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV
			ł			VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD
			ŀ			AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI
						EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ
			1			MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA
						VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ
		ļ				PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
		i				QHNTEIRMAVSKVADKMDHLMTKVEELQKH
						SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
			[LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
		l				NLMMEKRNNSLOTATENTOARVLHAEOEKA
						KVTEELAAATAQVSHLQLKMTAHQKKETEL
						QMQLTESLKETDLLRGQLTKVQAKLSELQET
				Ï		SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL
]			RVEKESLEKNLSERKKKSAQERSQAEEEIDEI
]	}]		RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS
	[.		LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
						EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA
						QITALTKQNEQHIKELEKNKSQMSGVEAAAS
						DPSEKVKKIMNQVFQSLRREFELEESYNGRTI
						LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE
						EKAEERPRRPSQEQSASASSGQPQAPLNRERP
			i i			ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR
						KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP
						TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN
						PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE
				,		GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
				İ		ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE
						DELFKGATLKALRPKAQPEEEDEDEVSMKGR
						PPPTPLFGDDDDDDDDDDWLG
817	2167	Α	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG
	·					

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, '=possible nucleotide deletion, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYTDNLWDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPG'HLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLTGLHACG\DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG QALSVLETEDS
819	2169	Α	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARNDMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\GRNEKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID SEQ ID Met SEQ Predicte	Predicted end Amino acid sequence (A=Alanine C=Cysteine,
NO: of NO: of hod ID NO: beginning	1
nucl- peptide in nucleoti	
eotide seq- USSN location	corresponding I=Isoleucine, K=Lysine, L=Leucine,
seq- uence 09/496 correspo	
uence 914 ng to fir	
amino a	
residue	
, , , , ,	
peptide	/=possible nucleotide deletion, \=possible
sequence	
	EMTNLKDIGLYNLRNITRG\AIRIEKNADLCYL
	STVDWSLILDAVSNNYIVGNKPPKECGDLCP
	GTMEEKPMCEKTTINNEYNYRCWTTNRCQK
	MCPSTCGKRACTENNECCHPECLGSCSAPDN
	DTACVACRHYYYAGVCVPACPPNTYRFEGW
	RCVDRDFCANILSAESSDSEGFVIHDGECMOE
	CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT
1	KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA
	SELENFMGLIEVVTGYVKIRHSHALVSLSFLK
	NLRLILGEEQLEGNYSFYVLDNQNLQQLWD
	WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE
	VTGTKGROSKGDINTRNNGERASCESDVLHF
	TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK
	EAPFKNVTEYDGQDACGSNSWNMVDVDLPP
	NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
	VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS
	NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP
	QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT
	ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE
	AEYRKVFENFLHNSIFVPRPERKRRDVMQVA
	NTTMSSRSRNTTAADTYNITDPEELETEYPFF
	ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE
	KLGCSASNFVFARTMPAEGADDIPGPVTWEP
	RPENSIFLKWPEPENPNGLILMYEIKYGSQVE
	DQRECVSRQEYRKYGGAKLNRLNPGNYTARI
	QATSLSGNGSWTDPVFFYVQAKRYENFIHLII
	ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN
	GVLYASVNPEYFSAADVYVPDEWEVAREKIT
	MSRELGQGSFGMVYEGVAKGVVKDEPETRV
	AIKTVNEAASMRERIEFLNEASVMKEFNCHH
	i
	VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR
	SLRPEMENNPVLAPPSLSKMIQMAGEIADGM
	AYLNANKFVHRDLAARNCMVAEDFTVKIGD
	FGMTRDIYETDYYRKGGKGLLPVRWMSPESL
	KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ
	GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL
	MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE
	VSFYYSEENKLPEPEELDLEPENMESVPLDPS
	ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF
	DERQPYAHMNGGRKNERALPLPQSSTC
821 2171 A 6691 106	825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE
	SNHVFCVSSMLRLQKRLASSVLRCGKKKVW
	LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT
	VHSRARCRKNTLARRKGRHMGIGKRKGTAN
, , , , ,	ARMPEKVTWMRRMRILRRLLRRYRES/KRYR
	ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH
	IHKLKADKARKKLLADQAEARRSKTKEARK
, , , , , ,	RREERLQAKKEEIIKTLSKEEETKK
822 2172 A 6715 772	
21/2 A 0/13 //2	
]	SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\
	SGDI/CNA/CVLL/LKRWKKLPAGSKK/NWNH
	VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST
	QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF
	TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\
	TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC
	SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823 2173 A 6727 3	4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG
	NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA
	INDICATIONAL PAGNACIAL CONTROL OF THE BALL
	SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

NO: of pripide sequence where the continuous process of the continuous process process of the continuous process of the continuous process process of the continuous process of the continuous process proce	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
seq- uence Sequence			hod				
uence 914 914 1				1			
uence 914 ng no first simio neid residue of peptide pepti	t			L	ľ		
amino seid residue of peptide sequence pictures with the sequence picture of peptide sequence per sequence pe	1 *	uciico	}	i .			
residue of peptide sequence Poptide uncleotide cleion, "possible nucleotide cleion," possible nucleotide cleion," possible nucleotide cleion," possible nucleotide cleion," possible nucleotide cleion," possible nucleotide cleion," possible nucleotide insertion SOPPSODOGESNYPSVGSLADPPYINTPOMN TEVPTLNSAAPASNSGAGVLPSPATPRESVPTP RTPTETPREFOGTASGGSSKYADSTDOGSS ASTESTTEPLNSVEPATMOPPERAHSLYVTLL. SOSVMINIFERENPSVEPATMOPPERAHSLYVTLL. SOSVMINIFERENPSVEPATMOPPERAHSLYVTLL. LOSVSNINKPERAPPOSCCACAMNINKGADVG LYPPOSSNEOPYRCTCGFSAIMNRKLGYNSGL LEDENGENDE CLACKTON, CANADIS CONTINUED C	1 001100			' ' '	, –		1 ` ' ' ' '
Peptide							
sequence Sequence Sequence		i					/=possible nucleotide deletion, \=possible
TEVTINSAAPASNSGGSVLYBSTATERSVPTE RTPETTPERGGTASGGSSVAYDSTDOGS ASTESTTEPI.NSVEPATMQPPEAHSLYVTLIL SDSWMIPKDRNPDSCCLACMMIKGADVG 1.YPDSSNEDQVRCTCGFSAIMMRKLGYNSGL 1.EDELDFGKNSDIGGAARRILMMQCSTFL PQVEGTKKPQEPPISLLILLQNGHTQPFASIN ILDVISSSNRROTIPEVSWSTDRVQADANDIY WTECFNALEQGRQYVDNFTGGKVDEALVRS ATVISWPHSNVLDISMLSSQDVVRMLLSLQP PLQDAIQKRTGRTWENIQHVQCPLTWQOFFI KMAGRGTYGSESPEPLPPTLUGYDKDFLT ISPS1.PFWERLLDPYGGHDWAYUVCPEN EALLEGAKTFRDLSAVYEMGRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWPOW SGEENDNHSRIKLYAQVCHHILAFYLATLQL DSSLLIPPY QTPPAAAQQATFCNAGAQQOQATCMAGPLAPN GSAAPPAGSAAPTISNSSSTNPAASSAASGSSSV PPVSSSAAFGISQISTTSSSGFGSVGQONST GGISADRTQGHIGCGGDTDPGGSSSQCPQDQ QESVTERERIGIPTEPDSASSHAIPPAVITYM VDPFTYAAEEDSTSGNFWLISJMRCYTEMLD NLPEHMRNSFLQIVPCGYMLQTMKDEQVFY IQVLSSMAFSVYQCCRRPLFTQHIKSLTGFGP AASIEMTIKNPEREPSIQLYSPFFLLAPIKDKQT ELGETTGGASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVIALEPMSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR CHIGHELWSILLGEGSLQTISKLKLKDVCRM CGISAADPSILSACLVAMEPQGSFVVMPDAV TIGSSFGRSTALNNQSQLNTPQDASCHILL VFTTSSTIQVAPANYPREDGSFVVMPDAV TIGSSFGRSTALNNQSQLNTPQDASCHILL VFTTSSTIQVAPANYPREDGSFSVMDDAFVDL LLGYFYSTAKEALBLOWWSSCPOAQMQC PLELKASLHHHISVAQTDELLPARNSGRYPHE LDSKTTSDLPVILEQVMALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALMMIL VFTSSTIQVAPANYPREDGGSFVAMPDAV TIGSSFGRSTALNNQSQLNTPQDASCHILL VFTTSSTIQVAPANYPREDGGSFVVMPDAV TIGSSFGRSTALNNQSQLNTPQDASCHILL VFTTSSTQVAPANYPREDGGSFVDAMDRQ CHALKASLHHHISVAQTDELLPARNSGRYPHP LDSKTTSDLPVILEQVANLSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFVVLTQLYNALSWLTCHPATQD RTSCLPVIJFV		İ			sequence		
RTPRITPRIPROGCT AS GOGS VK 7DS TDOGS P ASTEST TREINS PER TMO PREPABLICY TILL SDS VMMIPK DERNED SCIC ACAMANIK ADVOC 1 YPDOS SINE DOVER TO COF SAMINEK LEYNOGL 1 YPDOS SINE DOVER TO COF SAMINEK LEYNOGL 1 YPDOS SINE DOVER TO COF SAMINEK LEYNOGL 1 YPDOS SINE DOVER THE SAMINE SINE PROVEST KENDED PROTECT SAMINEK LEYNOGL 1 PO YESTIK KPOEPPISILLILLON, OHT OPPASI.N 1 FLOYIS SINING ILLE CON SINING ILLE COPY STATE AND THE PROVINCE TO THE PROVINCE ALVES A TYLE SYMPHON YES STATE AND THE PROVINCE ALVES 1 FLODA IQKKET GROWN YES YES STATE SAMINE							SSQPSQDGQESNVPSVGSLADPDYLNTPQMN
ASTESTTREINSVERATION/PERABLIS/VTLL SISSWAMPERDISPOSCICLEC/MINIKGADUG LYPDSSNEDQYRCTCGFSAIMNIKLGYNIGG ELDEIGENSDIGQAABERIAM/MOQSTEL PQVEGTKKPOEPPISILLILLQNOHTOPFASIL ELDYSSNINGTLEVSWSYDEV/ADANDIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY WTECSNALEGORQYVDNPTGGKVDAANIY RAAGAGTYGSESPELPPTLLVOYDKOFL RAAGAGTYGSESPELPPTLLVOYDKOFL RAAGAGTYGSESPELPPTLAAGGATFOOLS SALEJPKKTQTPAAAGGATFOOLSAYWOYDEN SGESNDMHSRALLYAQVCRIBLAPYLATLQL SSLLIPKKTQTPAAAGGATFOOLSAYWOYDEN GSAAPAGGSAFWYSNSSASGOSGORST GSAAPAGGSAFWYSNSSASGOSGORST GSAAPAGGSAFWYSNSSASGOSGORST GGISADRTQONIGGGGDTDPGOSSOGYSOD GSVTSTERERIGIFTEDSADSHAHPAVYTYM VOPPTYAAEDSTSGNPWLLSIMRCYTEMD NUPPHMNSSELQUYPCOYMACTMAEDEVEY QQYLKSMARSVYCQCRRPLPTQHHKSLTGGGF AASIEMTLKNPERSPIQLVSPPTHAPKDKOF ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVNIALPNRSRSKVSAR KIGLQKLWEWGIGVQMTSLPWRVVIGRLGC ELGETGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVNIALPNRSRSKVSAR KIGLQKLWEWGIGVQMTSLPWRVVIGRLGC CGISAADSSILSACLVAMEPGGSFVVMPDA TMGSWGGRSTALMMQSSQLITTPQASCCHLL VPTTSSTIQVAANYPNEGGFSPNDDMFVDL PFPDDMDDIGLAGGTGCTSCHLSRCAPEELKQQP LALGYFVSTAKAENLPOWFWSSCPANDASQC LALGYFVSTAKAENLPOWFWSSCPANDASQC LALGYFVSTAKAENLPOWFWSSCPANDASQC LALGYFVSTAKAENLPOWFWSSCPANDASQC PLFLKASLHHHISVAQTDELLPARNSQRVPIKM WTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VPTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVLQVALSWLTCNPATOD RTSCLPVIEVVLTQLYNAMMIL VFTTSSTDVLRRVCLQVALSWLTALLQHGAN VMGSISAADSSPYTKM KTFGFCALHAASQCHWKVQULLEAGADD MTTLESTTPLFLAYAQTGCKLSAISLIS SGANNVCQALDKATPLFLAQGGGGGGG GGGGTKRPGTTCPTAATSSTGGFGOMAAAAA KYGENDAACLGWRWTLEAGAGNTKCULLL LSGAGADPLLYCHOSWVJSTDVL GLAREGOVVLRRVCLLLQHGANAA	İ]				TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP
SISSYMNIFKDRNFDSCCICACNMIKGADYSGL I VIPDSSNEDQYRCTCGFSAIMNEKLGYNSGL FLEDELDIFGKNSDIQAAERIMMCQSTCH PLOYEGKRYSDIQAAERIMMCQSTCH POVEGTKKPQEPISILLLLQNGTOPFASLN FLDYISSNINGTLFCVSWSYDRVQADNNDY WTECFNALLGQRGVYDNPTGGKPASLV FLDYISSNINGTLFCVSWSYDRVQADNNDY WTECFNALLGGRGVYDNPTGGKPALVS ATVHSWPHSNVLDISMLSSQDVVRMLLSQP FLQDAIQKERTGRTWENIQHVQGPLTWQGPL KMAGRGTVGSESPEPLPIPILLYGVDKDFLT ISPFSLPFWERLLLDPYGGHEDVAYIVVCPBI BSPSLPFWERLLLDPYGGHEDVAYIVVCPBI SPFSLPFWERLLLDPYGGHEDVAYIVVCPBI SPFSLPFWERLLLDPYGGHEDVAYIVVCPBI SALEGAKTFFRDLSAYYEMCKLGQHKPICK VLRDGIMRVGKTVAQKLTIDELVSEWPNQW GGENDHSISLKLYAQVCRIHLAPTALTQL DSSLLIPPKYQTPPAAAQQATFGNAQPLAPP GSAAPPAGSAMPTSNSSSTMPASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGQONTST GGISADRTQGNIGGGGDTPGGSSSQPSQDG QESVTERERIGIFTEPDSADSHAHPPAVVITW VDPFTYAAEEDSTSORPWLISIMGCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY VQVLKSMAFSVVCQCRRPLTOHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPTLAPIKDKQT LIGHTFCEASQKYNLFVGVCLSHORWLL ASCTDLFIGELLETCVVNLALPNRSRSKVGAR KIGLQKLWEWCGIGVQMTSLPWGRWLL ASCTDLFIGELLETCVVNLALPNRSRSKVGAR KIGLQKLWEWCGIGVQMTSLPWGWLL ASCTDLFIGELLETCVVNLALPNRSRSKVGAR KIGLQKLWEWCGIGVQMTSLPWGWPMDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPFTSSTIQVATAVNTNEEQFSNNDDMFVDL PFPDDMDNDIGIMTOKNLSSSPNSVPSGSP SGIGVGSHFQISRSQGERLLSRAAPEELKQQP LALGYFVSTAKAENLPWFWSSFVPSGSP SGIGVGSHFQISRSQGERLLSRAAPEELKQQP LALGYFVSTAKAENLPWFWSSFVPSGNS SGIGVGSHFQISRSQGERRAPEELKQQP LILLSTEDYNATATSPCGGGGGGG GGGGTKRPGTGTYATSFSGEPONAASAP LSLLSPFSQGTTYATFSFGEPONAASAP LSLLSPFSQGTTYATFSFGEPONAASAP LSLLSPFSQGTTYATFSFGEPONAASAP LSLLSPFSGCALLAASGGPHKVQUILLAGANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGDWRLLLGRANDP NATTLEETTIPLFLAVENGGNOWLILLAGADPD NATTLEETTIPLFLAVENGGNOWLILLAGADPD NATTLEETTIPLFLAVENGGNOWLILLAGADPD NATTLEETTIPLFLAVENGGNOWLILLAGADPD NATTLEETTIPLFLAVENGGNOWLILLAGADPD NATTLEETTIPLFLAVENGGNOWLILLAGACADPD NATTLEETTIPLFLAVENGGNOWLILLGNAKAGAA KYGEWLPHLLVAGFDPLLLLCNSWIDSVISID LITLEFTNWKTLABQAGACLY	1			ļ			RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP
I.VPDSSNEOQYRCTCGFSAIMNRKLGYNGG	İ						ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL
FLEDELDIFGKNSDIGQAAERALMMCQSTEL POVEGTKKPQEPISILLLLLQNCHTQPFASLN FLDYISSNINGTILPCVSWSYDRVQADNNDY WTECTNALLEGGRQYVDNPTGGKVDALVSR ATVHSWPHSNVLDISMLSSQDVVRMLSLQP FLQDAIQKKRTGRTWENIQHVQCPLTWQQFF KNAGRGTVGSESSPEPLPIPILLVGVDKDFLT ISPSLPFWELLLDPYGGHEDVAYIVVCPBI RALLEGAKTFREDSAYYEMCKLQGHKPICK VLRDGIMRVGKTVAQKLTDELVSEWPNQW SGEENDHSRIKLVAQVCRHHLAPTVATLQL DSSLLIPPKYQTPPAAAQQATPGNAGPLAPN GSAAPPAGSAMPTSNSSSTMPASSSASGSSV PPVSSSASAPGISQISTTSSSGFOSQVGQNTSTS GGISARTQGNIGGGGDTDPGGSSQPSQDG QESVTERERIGIFTEPDSADSHAHPPAVVITW VDPFTYAAEEDSTSGNFWLLSIMRCYTEMLD NLPEHMRNSFILQIVPCQVMLQTMKDEQVFY VQVLKSMAFSVVCQCRRPLTOHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPTLAPIKDKQT LIGHTFGEASQKYNLVFVGYCLSHORWLL ASCTDLIFIGELLETCVVNLALPRSRSKVASG KIGLQKLWEWCIGVQMTSLPWVVIGRLGR LIGHGLKDWSILLGECSLOTISKKLKDVCRM ASCADLSTBLAGACVAMEPGGSTPATSRSSVPSPGSP SGIGVGSHFQINGTGNITAGRAFACT TMGSVFGRSTALNMQSSQLATPQDASCTHIL VPFTSSTQVATANYNTEDGFSPNDDMFVDL PFPDDMDDIGIGMTGNLSSSSPVSPGSP SGIGVGSHFQISRGQGGRAFGFQAMQAC CIGISAADSSILSAGLVAMEPGGSPNDDMFVDL PFPDDMDDIGIGMTGNLSSSSPVSPGSP SGIGVGSHFQISRGQGGRAFGCAMAGAC CIGISAADSSILSAGLVAMEPGGSPNDDMFVDL PFPDDMDDIGGMTGNLSSSSPVSPGSP SGIGVGSHFQISRGGGGRAFGFTGAMAGAC LIGHGLKDWSILLGECSLOTISKALDVCRM VGGSTACHLVGVMCTARPGDASCTHIL VFTTSSTQVATANYNTEDGFSPNDDMFVDL PFPDDMDDIGGMTGNLISSSSPVSPGSP SGIGVGSHFQGGRAFTPGAMAGAC CIGISAADSSILSAGLVAMPBOGSSPVSTGSP SGIGVGSHFQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	ŀ		}				1
PQVEGTKKPQEPPISLLILQNOHTOPFASIN FILDVISSINNRGIT PCVSWSYDRYQADANDDY WTECTNALEQGRQYYDNPTGGKVDEAL VRS ATVHSWPHSHVDLISMLSSQDYRMLLSLQP FILQDAIQKKRTGRTWENIQHVQGPLTWQQFH KMAGRGTYGSEESPRPIPITLLVGYDKDFLT ISPPSI,FFWERLLLDPYGGRIDVAYIVVCYEN EALLEGAKTFFRDLSAVYEMCRLGHRHCK VLRDGIMRVGKTVAQKLTDELYSEWFNQPW SGEENDNISRLKLYAQVCRHHLAPYLATLQL DSSLIPPKYQTIPPAAAQGGATNAQPLAPN GSAAPPAGSAFNTSINSSSTNPAASSSAGSSV PPVSSSASAPGISISTTSSGFOSOVGGQNFST GGSAAPRAGSAFNTSINSSSTNPAASSAGSSV PPVSSSASAPGISISTTSSGFOSOVGGQNFST GGGSADRTQGNICGGGDTDPGGSSOPSODG QESVYERERGIGITEPDSADSHAIPPAVVIYM VDPFTYAAEEDSTSGHFWLLSIMRCYTEMLD NIPEHMRNSFEL DIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQHIKSLTGFGP AASIEMTLKNPERPSPQLYSPALIPERAVVIGLGR LGHELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPGGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHL VFPTSSTQVARANYPREDGFSFNDDDMFVDL FPFDDMDDIGILMTGNILHSSPNSSYPSFGSS SGIGVSHFOHSSGGFELLSGAPSELKQQP LALGYFYSTAKAENLPQWFWSSCPQAQNOC PLFLKASIHHISVAQTDELLPARNSQRYPHP LDSKTTSDVLRFVLEQVYALSWLTCNPATQD RTSCLPHIFVVLTQLVNALMBL SGANDSGALASS QFFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAARGGNVKVLTLLKKGRSDDAASD QFFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAARGGNVKVLRLLKKGRSDDAASD NATTLEETTPLFLAVENGIDVLRLLQHGAN VNGSHSMCGWNSLHQASSGCHKKSULLLGKG WMPHEAAYHNSVECLQMLNADSSENYIKM KTEGFGCALHLAASQGHWKVLLLLGKGRSGHTKCVELL LSSGADPDLVCNEDSWOLPHIAAAQMGHTKI LDLIPLTNRACDTGLARSGOLASQLISAGSGLASD NATTLEETTPLFLAVENGQDDVLRLLLQHGAN VNGSHSMCGWNSLHQASSGCHTKCVELL LSSGADPDLVCNEDSWOLPHIAAAQMGHTKI LDLIPLTNRACDTGLARVSPVAAGONELHLAYCLKYEK FISHRYFLKGCSLGFWNHTYFFYNHALKQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTIVKLAPAVERMLSARASNAWIL LIFTLEFTIVKLAPAVERMLSARASNAWUL LIFTLEFTIVKLAPAVERMLSARASNAWUL LIFTLEFTIVKLAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWNIL LIFTLEFTIVELAPAVERMLSARASNAWIL LIFTLEFTIVELAPAVERM					'		
FLDYISSINRIĞTILPC/SWYDRYOGADNINDY WTECFNALE/OGRQYYDNPTOGRVDEAL VRS ATVHSWPHSINVLDISML SSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQOFH KMAGRGTYOSEESPEPLPITLVGYDKDFLT ISPFSLPFWERLLLDPYGGRRDVAYIVVCYEN EALLEGAKTFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWPNOPW SOEEDNINSRLK LYAQVCRHILAPYLATLQL DSSLLIPPKYQTPPAAAQGQATFGNAGPLAPN GSAAPPAGSAAPITSINSSITIPASSASAGSSV PPVSSSASAPGISQISTITSSSGFGGSVQGONET GGSAPPAGSAAPITSINSSTITASSASAGSSV PPVSSSASAPGISQISTITSSSGFGGSVQGONET GGSAPPAGSAAPITSINSSTITASSASAGSSV PPVSSSASAPGISQISTITSSSGFGSVSVGQONET GGSAPTGGRAFTYLLSLMRCYTEMLD NLPEHMRNSFLQIVPCQYMLLSLMRCYTEMLD NLPEHMRNSFLQIVPCQYMLLSLMRCYTEMLD NLPEHMRNSFLQIVPCQYMLQTMKDEQVEY QVYLKSMAFSVYQCGHLPTQHIKSLITGFGF AASIEMTLKNPERPSPIQLYSPFILAPIKDKQT ELGETTGEASQKYNVLFVQYCLSHDQRWLL ASCTDLHGELLETCVVNLAPHRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGALGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALINMQSSQLMTPQDASCTHIL VPFTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGLMTGNLHSPRNSSFVSPGOSP SGIGVGSHPPGHSRSQGERLLSREAPEELKQQP LALGYFYSTAKAENLRQWFWSSCPQAQNAC PLFLKASLHHIBVSAQTDELLFARNSQRVPHE DSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPHIFVVTQLYNALMSVLTCNPATQD RTSCLPHIFVVTQLYNALMSVLTCNPATQD RTSCLPHIFVVTQLYNALMSVLTCNPATQD RTSCLPHIFVVTQLYNALMSVLTCNPATQD RTSCLPHIFVVTQLPYALMSVLTCNPATQD ASLAYGRQLHGGGDRRGLLGRSSQDAASD QPFRCRSGSTAGRLVKQMDFTEATADTCSTV GLAAREGNVKVLRKLLKKGRSTDVADNRG WMPHEAAYHNSVECLQMLINADSSEBYIKM KTFEGFCALHLAASQGHWKVQILLEAGADP NATTLEETTPLLAVENOQIDVLRLLLIQHGAN VNGSISMCGWNSLHQASFGCHE DCLEILRNGYSPDAQACLVFOFSSPVCMAFQ KOCHSHMCGWSLHAASQGHKCVLLYFK SIGNAVNCQALDKATPLILAACSGGHKKVLKLYKK FISIFR YFLKGCSLGFWNHIYSVAFGGHE DCLEILRNGYSPDAQACLVFOFSSPVCMAFQ KDCEFGTOVILLKYGAGNELHLAYCLKYFK FISIFR YFLKGCSLGFWPHHYEFYNHALKQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSDIDT LIFTLFTINKALLAPAVERMLSARASNAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTFRINGARASANAWUL LILIPLTRINGARASANAWUL LILIPLTRINGARASANAWUL LILIPLTRINGARASANAWUL LILIPLTRIN	İ	ì	İ				
WTECTNALEGGRGY VDNPTGGK VDEAL VRS ATVHSWPHSHVDLISMLSSOPRMILLSLOP FLQDAIQKKRTGRTWENICHVOGPH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT ISPPSLPFWERLLLDPYGGRHDVAYIVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWRPQFW SGEENDINISRLKLYAQVCRHHLAPYLATLQL DSSLIPFKYOTPPAAAGGAPTGNAGPLAPN GSAAPAGSAFNTTSNSSSTNPAASSSAGSSV PPVSSSASAPGISQISTTSSSGPSGSGGGGGPNFST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESYTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEDSTSGNFWLIKKLYTEMLD NLPEHMRNSFLQIVPCQYMLQTMKDEQVEY IQVLKSMAFSVYCQCKRPLPTIKKSLTGFGF AASIEMTLKNPERPSPIQLYSPFILAPIKAKOT ELGETFGEASGKYNVLFVGYTSDORWLL ASCTDLHGELETCVVNIALPNSRSSKVSAR KIGLOKLWEWGGSVVTNIALPNSRSSKVSAR KIGLOKLWEWGGSVVTNIALPNSRSSKVSAR KIGLOKLWEWGGGVMTSLPGVTGAT CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNNQSSQLNTPQDASCTHIL VPFTSSTIQVAFANYPNEDGFSPNNDDMFVDL PPPDDMDNDIGILMTONLHSSSYVPSFGSS SGIGVGSHFGHSRSQGERLLSREAPEELKQOP LALGYFVSTAKAENLPQWWYNSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKSTTSDVLRFVLEQYNALATVCNPATQD RTSCLPVIFVVLTQLYNAMMIL 824 2174 A 6732 2440 365 VEEGLGRRTPFGGRGFVTFARGFDSVRR RLLPPSSAAAFSSHRNLLCSRRGGGGGG GGGGGTKRFGITGFTAATSPSGEPONAASAP LSLLSPPFQGTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGLLKQMPTFTAYADTCSTV GLAAREGNVKVLRLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLNADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLELAVENOQIDVLRLLQHGAN VNGSHSMCCWNSLHQASFGRHKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGTKIL LDLLIPLTNRACGSLGPWNHIVSPVSAVFGGHE DCLEILLRNQYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIVSPVSAVFGGHE DCLEILLRNQYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIVSPTFYNHALKAQA KYKEWLPHLJVAGFDPLILLCNSWTDSVSIDIT LIFTEFTIVELAPAVERMLSARASNAWIL LIFTLFFTNWALAPAVERMLSARASNAWAU	}		Į.	}			
ATVHSWPHSNVLDISMLSQDVYRMLLSLQP FLQDAIGKKR FRERWENICHVQGPLTWQOFF KMAGRGTYGSESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVA YIVVCPEN EALLEGAKTFRDLSAVYEMCRLQHKPICK VLRDGIMKVGKTVAQKLTDELVSEWPYNQFW SGEENDNISRLKLYAQVCRHHLAPYLATLQL DSSSLLPPKYQTPPAAAQGQATFGNAGPLAPN GSAAPPAGSAFNTSNSSSTPAASSSAGSSV PPVSSSASAFGISQISTITSSSGFSGSVGGQNFST GGSADRTQGNIGCGGDTDPGGSSCOPSDG QESVTERERIGITEPDSADSHAHPPAVYIYM VPPPTYAAEDSTSGHFWLLSKLYTMLDL NLPEHMRNSFLLQIVPCQYMLQTMKDEQVEY IQYLKSMAFSVYCQCRPLPTQIHKSLTGFGP AASIEMTLKNPERPSPIQLYSPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNLALPMSRRSSKVSAK KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVWPDAV TMGSVFGRSTALNMQSSQLMYPDDASCTHIL VPTTSSTIQVAPANYPNEDGFSPNNDDMFVDL PPPDDMDNDIGLIMTOHLHSSSVYSPGSS SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFYSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHSVAQTDELLFARNSQRVPHE LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPHIFVVTQLYNALMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPFGGRRGPVTFARRGPSSVRR RILLSPFSAAAFSSHIRHLLCERGGGGGG GGGGGTIKRPGITGTPTAATSPSGEPGNAASP LSLLSPFGQTTYQHPGVABFSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNKVLRRLLKKRSTDVADNIG WMPHEAAYHNSVECLQMLNADSSENYIKM KTFEGFCALHLAASQGHWKVQULEBAGADP NATTLEETTPL-LAVENQQIDVLRLLIQHGAN VNGSISMCGWNSLHQASFGPHKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLSGGADPDLYCNEDSWQLPHAAAQMGHKIL LLLIPLTRRACDTGLAKTPVSAPGGHE DCLEILRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILKYGAQNELHLAYCLKYEK FISIFRYFLKCGSLGPWNHYFFYNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSUDIT LIFTLFTINKALAPAVERMLKARASNAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LILIPLTFRINKALARASANAWHIL LIFTLFTIPLKLARASANAWHIL LIFTLFTINKALARASANAWHIL LIFTLFTINKALARASANAWHIL LIFTLFTINKALARASANAWHIL LIFTLFTINKALARASANAWHIL LIFTLFTINKALARASANAWHI	1						1
FILODAIQKKRTIGRTWENIQHVOGETLWQCPH KMAGRGTYGSESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVA YIVVCPEN EALLEGAKTIFTDLSAVYEMCRLQOHKPICK VIRDGIMRVGKTVAQKCHHLLAYLATLQL DSSLLIPPKYQTIPPAAAQQATTGMAGPLAFN GSAAPPAGSAFNIPTSNSSTNIPAASSASGSSV PPVSSSASAFOIGSISTTSSSSSSGSSVGOGNET GGISADRTQGNIGGGGDTDPGQSSSSJGGONET GGISADRTQGNIGGGGDTDPGQSSSSOSGOOD QESVTERERIGHTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NILPEHMRNSFILQIVPCQVMLQTMKDEQVFY IQYLKSMAFSVYQCCRREPLTQHHKSLTGFGF AASIEMTLKNPERPSPIQLYSPFFLLAKDGOT ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTEPWVVIGTLGR LGHGELKDWSILLGEGSLQTISKKLKDVCRM CGISAADSFSILSACLVAMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFITSSTIQVATANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSPNSSPYPSPGSS SGIGVGSHFQHSRSQGERLLSREAPELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSRQFPHP LDSKTTSDVLRRVLEQVALSWLTCNPATQD RTSCLPVIHVVLTQLYNAIMMIL 824 2174 A 6732 2440 365 VEEGLGRRTPPGGRRGPVTPARPGPDSVR RLLSPSFAGGTTAKAENLPQFTSVR RLLSPSFAGGTTYQHGVALSWLTCNPATQD RTSCLPVIHVVLTQLYNAIMMIL SELLSPFFGGTTYQHGVALSWLTCNPATQD RTSCLPVIHVVLTQLYNAIMMILLSRRRGGGGGG GGGGGTIKKPGTTTYATSPGGFRMASSAP LSLLSPFFGGTTYQHGVALSWLTCNPATQS QPFRCRSGSTAGRLVKQMACHSATSTAGNASSAP LSLLSPFFGGTTYQHGVALSWLTCNPATQSTV GLAAREGNVKVLKKLLKKGRSVDVAQDNG WMPHEAAVHINSVECLQMANADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGGIDVLRLLLQHGAN VNGSHSMCGCMSNSLHQASFQENAEIRKLLLRG GANKECQDDFGITFLFVAAQTGKLLGSSGGHE CGANNCQALDKATPLFTAAQGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGANNCQALDKATPLFTAAQGGGHTCCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSSGADPDLYCNEDSWQLPHAAQAGGHTKCVELL LSS				ĺ			1
KMÄGRĞTYGSEESPEPLPİPTILÜGYDKÖPLT ISPFSLIFWERLLIDPYGATIVAVIVOLET ISPFSLIFWERLLIDPYGATIVAVIVOLET BALLEGAKTFFDLSAVYEMCRLGQHKPICK VIRDGIMRVGKTVAQKLTDELVSEWINQPW SGEENDINISRIKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQIFPAAAQQATIGMAGPLAFN GSAAPPAGSAFNPIYSISSTSIPPAASSSASGSS PPVSSSASAPGISQISTISSSGFGSVGQNFST GGISADPTGGNIGGGGDTPQGSSSOPSODG QESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSLIMKÇTIEMLD NILPEHMRISPERIQIPCYQMLQTMKDEQVEY IQYLKSMAFSVYCQCRRPLTQHIKSLTGFGF AASIEMTLKNPERSPIQLYSPFILAPIKDKQT ELGETGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGBLJETCVVNLAPSRRSK VSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CIGISAADPSSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHLL VFPTSSTIQVATANYNEGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPYSPGSS SGIGVOSIFPGINSROGERLJSREAPEELKQOP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTIDELLPARNSRVPFP LDSKTTISDVLRVLEQVNALSWLTCNPATQD RTSCLPVIIFVVLTQLVALSWLTCNPATQD RTSCLPVIIFVVLTQLVALSWLTCNPATQD RTSCLPVIIFVVLTQLVALASWLTCNPATQD RTSCLPVIIFVVLTQLVALAMMIL 824 2174 A 6732 2440 365 VEEGLGRRTPFGGRGRVFVTARPGFDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTIKRPGGRRFVTPARPGFDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTKRPTPGGRGFVTPARPGFDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTKRPTPGGRGFQVTYARPGFDSVVR KTFGGFCALHLAASQGHWKIVQLLLGAADP NATTLEETTIPLFLAVERGDDVLRLLLOHGAN VNGSHSMCGWNSLHQASFQENAEIKLLLRK GANKEQQDDFGITLFVAAQYGKLLEAGADP NATTLEETTIPLFLAVERGDDVLRLLLOHGAN VNGSHSMCGWNSLHQASFQENAEIKLLLRK GANKEQQDDFGITLFVAAQYGKLEAGADP NATTLEETTIPLFLAVERGDDVLRLLOHGAN VNGSHSMCGWNSLHQASFQENAEIKLLLRK GANKEQQDDFGITLFVAAQYGRLEAGADP NATTLEETTIPLFLAVERGDDVLRLLLOHGAN VNGSHSMCGWNSLHQASFQENAEIKLLRK GANKEQQDDFGITLFVAAQYGKLEAGADP NATTLEETTIPLFLAVERGDDVLRLLLOHGAN VNGSHSMCGWNSLHQASFQENAEIKLLLRK GANKEQQDDFGITLFVAAQYGCHELLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDFLILLCNSWIDSVSIDT LIFTLEFTIWKTLLAVACKERMLSARASNAWIL LIFTLEFTIWKTLLAVACKERMLSARASNAWIL							
ISPSI.PFWERLLLDPYGGHRDVA YIVVCYEN EALLEGAKTFRDLSA YYEMCRI.GQHKPICK VIRADGIMR VGKTVAQKITDELVSEWFNQPW SGEENINHSRI.KLYAQVCRIHILAPLATLQL DSSLLIPPKYQTIPPAAAQQATTGNAGPLAPN GSAAPPAGSAFNPTSNSSTNPAASSAGSSY PPVSSSASAFOISQISTTSSGSGSVGQQHPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVIYTM VDPFTYAAEDSTSGNIPWLISIJMRCYTEMLD NILPEHMRNSEILQIVPCQVMLQTMKDEQVFY IQVLSSMASFVVQCQREPLETQHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPEILAPIKDKOT ELGETFGEASQKYNVLPGVGUSHDQRWLL ASCTDLHGELLETCVVNIALPPRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPRRVYQGRUG CGISAADSPSILSACLVAMEPQGSFVMPDAV TMGSVFGRSTALNMQSSQLTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVMPDAV TMGSVFGRSTALNMQSSQLTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVMPDAV TMGSVFGRSTALNMQSSQLTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVMPDAV TMGSVFGRSTALNMQSVGJORFQNASOCHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL VFPTSSTIQVAPANYPNEDGFSPNNSPVPSGSP SGIGVGSHFQHSRSQGERLISREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTIDELIPARNSQRVPHP LDSKTTSDVLRFVLEQVNALSWLTCMPATOD RTSCLPVHIFVVLTOLYNALMMIL VFPTSSTQAAASAP LSLSPFGGTTYQHFVARSQRGFGGGGG GGGGGTTKKPGGTPTARTSGPPDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTTKRPGGTTYQHGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKGGRSGDAASD QFFRCRSGTAGRLVKQMDFTEAYADTCSTV GLAARGGNVKVLRKLKGRSVDVADNRG WMPHEAAVHNSVECLQMINADSSENYIKM KTEGGCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGDDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIRLLLRK GANKECQDDFGITPLFVAAQYGKLESLVSILIS SGANVNCQALDKATPLFIAAQEGTIKCVELL LSSGADPDLYCNEDSWQLPHAAAQQHTIKL LDLLIPLTINRACOTIGLINKYSPYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KVEWLPHLLVAGFDPLILLCNSWDVSDYSAVFGGHE DCLEILRNGYSPDAQACLVFGFSSPVCMAFQ KVEWLPHLLVAGFDPLILLCNSWDVSDYSAVFGGHE DCLEILRNGYSPDAQACLVFGFSSPVCMAFQ KVEWLPHLLVAGFDPLILLCNSWDSVSDT KVEWLPHLLVAGFDPLILLCNSWDSVSDT KVEWLPHLLVAGFDPLILLCNSWDSVSDT LTFLEFTIVKYTLLAPVERKTLAPACKYSPY SAVFGGHE DCLEILRNGYSPDAQACLVFGFSSDAASSNAWIL LIFLEFTIVK	j		ļ	1			
EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLEVLSEWFNOPW SGEENDNHSRLKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQTPPAAAQQQATIGNAGPLAPN GSAAPPAGSASPNPTSNSSSTSNPAASSSASGSSV PPVSSSASAPOISQISTTSSGFGGSVGQNFST GGISADPTGQNIGGGGDTDFQGSSSOPSQDG QESVTERERIQIPTEPDSADSHAHPPAVVTYM VDPPTVAAEDSTSGNFWLSIMECYTEMLD NLPEHMRNSFILQIVPCQVMLQTMKDEQVFY IQYLKSMAFSVYCQCRRLPTQHHKSLTGFG AASIEMTLKNPERSPIQLYSPPFILAPIKDKOT ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGGLLETCVVNLGYSPPFILAPIKDKOT ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGGLLETCVVNLGSPPFILAPIKDKOT CGISAADSPSILSACLVAMEPQGSFVVMPPAV TMGSVFGRSTALNMQSSQLNTPQDASCTHLL VPFTSSTIQVATANYNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPYSFGGS SGIGVGSIFFORISRGGGETLSREAPEELKQOP LALGYFVSTAKAENLPQWFWSSCPDAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQVNALSWLTCNPATQD RTSCLPVHFVVLTQLYNALMMIL 824 2174 A 6732 2440 365 VEEGLGRRRTPFPGGRGFQFVTPARPGFDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGTKRFPGGRGFQFVTPARPGFDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTKRPGICHANASSP LSLSPPFGOTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD PLSLSPPFGGTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSGDAASD PNATTLEETTPLFLAVENGDWLRLLLGAGADP NATTLEETTPLFLAVENGDWLRLLLGAGADP NATTLEETTPLFLAVENGDWLRLLLGAGADP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLGRAGDP NATTLEETTPLFLAVENGDWLRLLLCHK GANKECQDDFGTITLFVAAQFGKLLGRAGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFLAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVELL LSSGANVNCQALDKATPLFAAQGGHTKCVEL							
SGEENDNHSRI, KLYAQVCRIHILAPYLATLOL DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNITSNSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGTSGSVGQONTST GGISADRYQGNICGGDTDFQGSSSQFSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFVLISLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQVLKSMAFSVYQCCRPPLPTQHHIKSLTGFGF AASIEMTIKNPERPSPLYSPPHLAPIKDKOT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLMTPQDASCTHIL VFPTSSTIQVAPANYPNEDGIFSPNNDDMFVDL PFPDDMDNDIGILMTGINLHSSPNSSPVPSFGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPWWSSCCPAQONQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPQFDSVRR RLLPPSSAAAFSSIRKHLCSRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSFFQGTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYDDTCSTV GLAAREGNVKURKLLKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLESTTPLFLAVENOGIDVURLLLQHGAN VNGSHSMCGWNSLHQASFQERABIIKLLLIKR GANKECODFGITPLFVAAQYGKLESISILS SGANNNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LLSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSVSVSAVFGGHE DCLEILLRRNGYSPDAQACLVGFGSSPVCMAFQ KDCEFFGIVNIILLKYQACELHLAYCLKYEK FSIFRYFLRKGCSLGPWMHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARARSNAWIL						:	
DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSGFSQSVGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERGIGPTEPISADSIAHPPAVVIYM VDPFTYAAEEDSTGNFWLLSIMRCYTEMLD NLPEHMRNSFILQIPCQYMLQTMKDEQWPY IQYLKSMAFSVYCQCRRPLPTQHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPPRSRSKSVSAR KIGLQKLWEWCIGIVORLPWRSVRGRIGR LGHGELKDWSILLGCSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHLL VFFTSSTIQVAFANYPEDGFSPNDDMFVDL PFPDMDDIGILMTGNLHSSPNSSVPYSPGSPS SGIGVGSHPOHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRTTPFGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTTKRFGITGFTAATSPSGEPGNAASAP LSLLSFFGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRQHRGGDRKGLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYDDTCSTV GLAAREGNVKURKLKGRSVDVADNRG WMPHEAAYHNSVECLQMLNADSSENYIKM KTFEGFCALHLAASQGHWKIVQULLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMGGWNSLHQASFGENAEIIKLLLKR, GANECQDDFGITPLFVAAQYGKLESISJIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGAPPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTINRACDTGLINKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCFFFGIVNILLKYGADCHELHAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWDSVSIDIT LIFTLEFTNWKTLAPAVFRMLSARARNAWIL]				VLRDGIMRVGKTVAQKLTDELVSEWFNQPW
GSAAPPAGSÁFNPTSNSSSTNPAASSSASGSSV PPVSSSASAFGIQISTTSSGFSGGVGGQNPST GGISADRTQGNIGCGDTDPGQSSSQPSQDG QESVTERERGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEDDSTSMVILLSIMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVPY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASLEMTLKNPERRSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVNILAPNRSRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPTTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHPQHSRSQGERLLFARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPFGGRGFVTPTARPGPDSVRR RLLPPSSAAAFSSHRINLLCSRRRGGGGGG GGGGGTIKRPGTGTPTAATSPSGEPGNAASAP LSLLSPPGGTTYQHVAVENSCGPQBVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPTRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLKKLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLNADSSENYIKM KTFEGFCALHLAASQHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIKLLLRK GANKECQDDFGITPLAVAQYGKLESLISILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVVSAVFGGHE DCLEILRNQYSPDAQACLVFGFSSPVCMAFQ KDCEFGGTVNILLKYGQNNEHLIAVCLKYEK FSIFRYFLRKGCSLGPWNHYEFVNHAIKAQA KYKEWLPHLLVAGGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAKGASNAWIL							SGEENDNHSRLKLYAQVCRHHLAPYLATLQL
PPVSSSASAPGISQISTTSSSGFSGSVGGONPST GGISADRTQGNIGCGGDTDPQQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSIMRCYTEMILD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASIEMTLKNPERRSPIQLYSPFILLAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPRISRRSK VSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMSQSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNHLSSPNSSVPSPGSP SGIGVGSHFQHRSSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVADTELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAMMIL 824 2174 A 6732 2440 365 VEGIGRRTPFPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRIHLLCSRRGGGGGG GGGGTIKPGITGFTAATSPSGEPONAASAP LSLLSPFSQGTTVPGRGGRGPVTARPGPDSVRR RLLPPSSAAAFSSHRIHLLCSRRGGGGGGG GGGGGTIKPGITGFTAATSPSGEPONAASAP LSLLSPFFQQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKGLLGRSSGDAASD QFFRCSGSTAGRLVKOMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFGGFCALHLAASQMWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLQLAGFQENAEIIKLLLRK GANKECQDDFGITFLFVAAQYGKLESLISILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPLYCNEDSVQHAHAAQAGMTKIL LDLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEBLLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGTOWILLKYGAQNINEHLATVCLKYEK FSIFRYFLRKGCSLGPWNHYFFVNHAIKAQA KYKEWLPHLLVAGGTDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN
GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERRIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSORFWILSILMRCYTEMLD NLPEHMRNSFILQIVPCYMUQTMKDEQVPY iQVILKSMAFSVYCORPLPTOHINISLITGFGF AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRSKVSAR KIGLQKLWEWCIGIVQMTSLEWRVVIGRLGR LGHGELKDWSILLGSCLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPFTSSTIQVAFANYPNEDGFSPNNDDMFVDL PFPDDMDNIGILMTGNILBSSPNSSPVPSPGSP SGIGVGSHPGHSSQGERLLSREAPEELKQOP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLFARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNALSWLTCNPATQD RTSCLPVIIFVAQQCORLUNALDLEAGADP NATTLEETTPLFLAVENQQIDVLRLLLQHGAN VNGSHSMCGWNSULQLAGQDFONLICLLCNSWTDSVSDT LDLIPLINRACDTGLNKVSPVVSAVFGGHE DCLEILRNQYSPDAQACLVFGSSSPVCMAFQ KDCEFFGGTVNILLKYPTVNALAQAG KYCEMPTQUNALSWLTCNSWTDSVSDT LIFTLEFTNWKTLARASNAWIL LIFTLFFTNWKTLARASNAWIL LIFTLEFTNWKTLARASNAWIL							•
OESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYQQCRRPLFTQHIKSLTGFGF AASIEMTLKNPERPSPIQLYSPPFILAPIDKQCT ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELETCVVNIALPNESRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTAKIMQSSQLNTPQDASCTHIL VFFTSSTIQVAFANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGITMGTNIHSSPNSSPVSPGSF SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLFARNSQRVPHP LDSKTTSDVLRFVLEQVNALSWLTCNPATQD RTSCLPVIHFVLTQLYNAIMMIL RTSCLPVIHFVLTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMMIL STSCLPVIHFVTQLYNAIMM	ŀ						
VOPFTYAAEDSTSGNFWLLSLMRCYTEML.D NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYQCCRRPLPTQHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETTGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILGECSLQTISKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPPTSSTIQVAPANYPNEDGFSPNDDMFVVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSGSP SGIGVGSIFQHSRSQGERLISREAPELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAMMIL 824 2174 A 6732 2440 365 VEEGLGRRTPPGGRGFVTPARRGFDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDKRGLLGRSSGDAASD QFFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQDVLRLLLQHGAN VNGSHMCGWNSLHQASFGENBEIKKLLIRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SGANVNCQALDKATPLFIAQGEGHTKCVELL LSSGAPDDLYCHEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNNELHLAYCLKYEK FSIFRYPLRGCSLGPWNIVSFVYNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEETTNWKTLAPAVERMLSARASNASWIL		1					
NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHKISLTGFGP AASIEMTILKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPRRSRRSKYSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPFTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNILHSSPNSSPVPSPGSP SGIGVGSHFQHSSRQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRGGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD OPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENQQIVLRLLLQHGAN VNGSHSMCGWNSLHQASPGPENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLISILIS SGVANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCHEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRRGCSLGPWNIIVSFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL		1					-
IQYLKSMAFSVY.QCRRPLPTQHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETTGEASQKYNVLFVQYCLSHDQRWLL ASCTDLHGELLETCVVNIALPIKSRRSK VSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPFTSSTIQVAFANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSFNSSPVSSPGSP SGIGVGSHPGNSSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL S24 2174 A 6732 2440 365 VEEGLGRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPONAASAP LSLLSPFGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFGRLQHFUSVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGRDVAC ASLVFRACHTPROVERSAYGGGRDVAC ASLVFRACHTPROVERSAYGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		ĺ					I
AÁSIEMTLKNPERPSPIQLVSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALINMQSSQLNTPQDASCTHIL VPFTSSTIQVAFANYPNEDGISPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSFNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRYPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRTPFGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFFGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAARGGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTIPLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITELFVAAQYGKLESLISILIS SGANVNCQALDKATPLFIAAQGGHTKVCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIRYYLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	1						
ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKYSAR KIGLQKL WECIGIVQMTSLPWR VVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLMTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMMIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPSSAAAFSSIHNILLCSRRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPQQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGVSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	1						
ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGHLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLLQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCNPATQD RTSCLPVILQYNALSWLTCLKPRATQD GGGGGTIKRPGTGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLNADSSENYIKM KTFEGFCALHLAASQGHWKIVQLLLEAGADP NATTLEETTPLFLAVENGQUDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLVAAQYGKLESL\SILIS SGGANVNCQALDKATPLFIAQGEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEBLLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILKYGAQNIELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAFANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNILHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPFGGRRGPVTPARPGPDSVRR RLLPPSSAAFSSHRIHNLLCSRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SGANNVCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEELLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNNELHLAYCLKYEK FSIFRYFLRKGCSLGFWNHIYEFVNHAIKAQA KYKEWPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							,
CGISAADSPSILSACLVAMÉPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTIKRPGHTGPTAATSPSGEPGNAASAP LSLLSFPFQGTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QFFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNSEHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWPHLLVAGFDPILILCNSWTDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							1
TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRINLLCSRRRGGGGGG GGGGTTKRPGHTGATASPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGIANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLK YEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWPHILLVAGFDPILILLCNSWTDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							LGHGELKDWSILLGECSLQTISKKLKDVCRM
VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSFGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSFFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QFFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGVANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							CGISAADSPSILSACLVAMEPQGSFVVMPDAV
PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHPGHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRINLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SGGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHILVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRINLLCSRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGQQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	}						
LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRINILLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLISILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVIIFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSFPFQQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQIFRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLISILIS SGVANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
RTSCLPVHFVVLTQLYNAIMNIL 824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNILCSRRRGGGGGG GGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\(\text{ANVNCQALDKATPLFIAAQEGHTKCVELL}\) LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL		j .					, ,
824 2174 A 6732 2440 365 VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYG\KLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	824	2174	A	6732	2440	365	
GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	1				- 1 -		
ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	f i					İ	LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC
GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							- 1
WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							-
KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	[
NATTLEETTPLFLAVÈNGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLRK GANKECQDDFGITPLFVAAQYG\KLESL\SILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL			.				3
VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	ļ .						` '
GANKECQDDFGITPLFVAAQYG\KLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
SGIANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL	!						
LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL					ĺ		
LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL						İ	
KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL							
LIFTLEFTNWKTLAPAVERMLSARASNAWIL						Ì	•
QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS]						
	لــــا		<u> </u>				QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G
825	2175	A	6735	277	1252	RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARA\TDYLQ ASAITRIPSYRYRYQRRSRSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	2176		6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IIHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS HHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VGKGIMGVFTKPIGGAAELVSQTGYGILHGA GLSQLPKQRHQPSDVHADQAPNSHVKYVW KMLQSLGRPEVHMALDVVLVRGSGQEHEGC LLLTSEVLFVVSVSEDTQQQAFPVTEIDCAQD SKQNNLLTVQLKQPRVACDVEVDGVRERLSE QQYNRLVDYITKTSCHLAPSCSSMQIPCPVVA AEPPPSTVKTYHYLVDPHFAQVFLSKFTMVK NKALRKGFP
827	2177	A	6748	2	1662	FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK ASGVSPTLWRPQAAATGLEMPSSGRALLDSP LDSGSLTSLDSSVFCSEGEGEPLALGDCFTVN VGGSRFVLSQQALSCFPHTRLGKLAVVVASY RRPGALAAVPSPLELCDDANPVDNEYFFDRS SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DTEDQESQHESEQDFSQGPCPTVRQKLWNIL EKPGSSTAARIFGVISIIFVGVSIINMALMSAEL SWLDLQLLEILEYVCISWFTGEFVLRFLCVRD RCRFLRKVPNIIDLLAILPFYITLLVESLSG\SQT TQELENVGAHCPGCLRLLRALVMLKAWGR HSTGLRSLGMTITQCYEEVGLLLFLSVGISIF STVEYFAEQSIPDTTTGKIVAFMCILSGILVLALPI AINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF
828	2178	,	6786	5672	1360	GTHPASSGPVPLPPAAVSAATREELGEPVFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPIPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNIKAPLTMASPAMLGNVESG GPPPTTASQPASVNIPG\SLPSSTPYTMPPEPTL SQNPLSIMMSR\MSKFAMPS\SNPGYNHDAI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	f	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
]		1]	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
<u> </u>		l		sequence		nucleotide insertion
			 	sequence		KTVASSDDDSPPARSPNLPSMNNMPGMGINT
l		1		·		QNPRISGPNPVVPMPILSPMGMTQPLSHSNQ
1		ĺ	1	•		MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH
i i						
						GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ
		ł				,
1		ŀ				SSADAALCKPGGPGGPDSFTVLGNSMPSVFT
j j						DPDLQEVIRPGATGIPEFDLSRIPSEKPSQTLQ
						YFPRGEVPGRKQPQGPGFGFSHMQGMMGEQ
						APRMGLALPGMGGPGPVGTPDIPLGTAPSMP
						GHNPMRPPAFLQQGMMGPHHRMMSPAQST
						MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
						HPGPVGSPGMMMSMQGMMGP\NRTS
829	2179	Α	6797	433	3	ASFFNFSICICKIÏLEVGPPVGHPAHDDVGGRH
			i			GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP
						ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG
						AAASSTSSGASCPPVPASSRWGVRSRTRSGSG
i l						GEREPRDRPSERPRLV
830	2180	A	6800	3	1911	LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL
						DREPRAPGPWLCPSRAGTAQDPARIRERRGR
						VAGGAAGPAMELRARGWWLLCAAAALVAC
						ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ
j j						AEISGEHLRICPOGYTCCTSEMEENLANRSHA
						ELETALRDSSRVLQAMLATQLRSFDDHFQHL
						LNDSERTLQATFPGAFGELYTQNARAFRDLY
[SELRLYYRGANLHLEETLAEFWARLLERLFK
						QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\
	'					RELRLRAT\RA\FVAAR\SFVQGLGVAS\DVVR
i						KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP
						GARPCPDYCRNVLKGCLANQADLDAEWRNL
						LDSMVLITDKFWGTSGVESVIGSVHTWLAEA
						INALQDNRDTLTAKVIQGCGNPKVNPQGPGP
}						EEKRRRGKLAPRERPPSGTLEKLVSEAKAQL
}						RDVQDFWISLPGTLCSEKMALSTASDDRCWN
						GMARGRYLPEVMGDGLANQINNPEVEVDIT
						KPDMTIROOIMOLKIMTNRLRSAYNGNDVDF
						QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS
						SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL
						LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA
021	2101	А	0000	ــــــــــــــــــــــــــــــــــــــ	1344	EGRLREKLFSGYDSSVRPAREVGDRVRVSVG
						LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS
						WDPAEHDGIDSLRITAESVWLPDVVLLNNND
						GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS
						IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT
						GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS
						RLIQPPGDPRGGREGQRQEVIFYLIRRKPLFY
						LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF
						ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT
				1		MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV
[[RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP
						GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL
					İ	SAPDLRRFIDGPNRAVALLPELREVVSSISYIA
						RQLQEQEDHDALKEDWQFVAMVVDRLFLW
	(TFIIFTSVGTL\VIFLDATYHLPPPDPFP
832	2100	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK
	2182					ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY
	2182	••		1		
	2182					=
	2182					DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN
	2182					DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
	2182					DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV
833	2183	Α	6846	116	602	IMPCNWWVARMLGRV EAEGEQVCGAKCCGDAPHVENREEETARIGP GVMESKFERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVVI LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVPPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6855	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL
837	2187	A	6863	2	1615	PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEEFKVL LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PNPRLEKIPVILSEFPTKIPRGELDMPVENGLR APVSAYQYALANGDVWKVHEVPDYSMAYG NPGVADATPPWSSYKEQSPQTLLELKRQRAA AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI EEMEEKVHGCCRIS
838	2188	Ā	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS LRPRKLDFFRSEKELNHLAVDEASGVVYLGA VNALYQLDAKLQLEQQVATGPVLDNKKCTP PIEASQCHEAEMTDNVNQLLLVDPPRKRLVE CGQLLKGNCALRALSNISLRLFYEDGSGEKSF VASNDEGVATVGLVSSTGPGGDRVLFVGKG NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQQFVAAFEDGPYVFFVFNQD KHPARNRTLLARMCREDPNYYSYLEMDLQC RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK SFPCGSEHLPYPLGSRDGLIRGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCFGESPPHPARVEGEAVICNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPRRPVVEQALYQFSNLLNSKSFLINFHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLLLSKVGVSQQPEDSQQDLPGERHALL EEERRVWHLVRPTDEVDEGKSKRGSVKEKE RTKAITEIYLTRLLSVKGTLQGFVDNFFQSVL APGHAVPPAVKYFDFLDEQAEKHNIQDEDTI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- ucnce	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL YAKEISTYKKMVEDYYKGIRQMVQVSDQDM NTHLAEISRAHTDSLNTLVALHQLYQYTQKY YDEJINALEEDPAAQKMQLAFRLQQIAAALE NKVTDL
839	2189	A	6872	1	1485	RARRIALQCHVCVCALTPGEQSGRRLPGQT WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH EDQTDCSSLRDENNKENYPDAGALVEEHAPP SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP KSIFKAESGRSHGESQETEHVVSSQSECQVRA GTPAHESPQNNAFKCQETVRL\QPRIDQRTAT SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS TQSVLA\DGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLLCTDV
840	2190	Α	.6873	2		FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN YLKQPVVKEKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTNLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSNDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVNSIKPSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA
						VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAJIRQARQRREKRKEYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR 'QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHE\APSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192	A	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSR\KTRHNST\DLPL
843	2193	Α	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195	Α	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
846	2196	A	6944	42	2672	RRKMAGCRGSLCCCCRWCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP YLEVLETMDNKKGRRYEAVKWMVVFAIGV CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE AGSGITEGKCYLYARQVPGLVRLPTLLWKAL GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL TWKVLFCSMSATTFLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLLGATFNCLNKRLAKYRMRNVHP KPKLVRVLESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGIITRHNLTYEFLQARLRQHYQTI
847	2197	A	6951		1994	NTINSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2199	A	6999	963		LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GIFYGTAIIMYVGPRYGNPKEQKKYLLLFHS LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
				1		DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG KVARRRVGATWLLHLAVADLLCCLSLPILAV PIARGGHWPYGAVGCRALPSIILLTMYASVLL LAALSADLCFLALGPAWCLRFS/GACGVQVA CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA SCHSALLCWAARRCRPLGTAIVVGFFVCWAP YHLLGLVLTVAAPNSALLARALRAEPLIVGL ALAHSCLNPMLFLYFGRAQLRRSLPAACHW ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	Α	7011		2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI SRGVLVCDECCSVHRSLGRHISIVKHLRHSA WPPTLLQMVHTLASNGANSIWEHSLLDPAQV QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG QTLQAELLVVYGADPGSPDVNGRTPIDYARQ AGHHELAERLVECQYELTDRLAFYLCGRKPD HKNGHYIIPQMADSLDLSELAKAAKKKLQAL SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD NLELSLRSQSDLDDQHDYDSVASDEDTDQEP LRSTGATRSNRARSMDSSDLSDGAVTLQEYL ELKKALATSEAKVQQLMKVNSSLSDELRRLQ REIHKLQAENLQLRQPPGPVPTPPLPSERAEH TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL YRIKGVSASAVPFTPSSPLLSCSQEGSRHTSK LSRHGSGADSDYENTQSGDPLGLEGKRFLE LGKEEDFHPELESLDGDLDPGLPSTEDVILKT EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA KAAKQLVTITTREKKQ
852	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL TVKGLLKPSFSPRNYKALSEVQGWKQRMAA KELARQNMDLGFKLLKKLAFYNPGRNIFLSP LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID QRLQPQRKFLEDAKNFYSAETILTNFQNLEM AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM DERGTEGAAGTGAQTLPMETPLVVKIDKPYL LLIYSEKIPSVLFLGKIVNPIGK
853	2203	Α	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV ARHVAAGAGHENKHGGSRRFPAGVAPRRAM ANVSKKVSWSGRDRDDEEAAPLLRRTARPG GGTPLLNGAGPGAARQSPRSALFRVGHMSSV ELDDELLEPLDMDPPHPFPKEIPHNEKLLSLKY ESLDYDNSENQLFLEEERRINHTAFRTVEIKR WVICALIGILTGLVACFIDIVVENLAGLKYRVI KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI AMGVVGGVLGAVFNALNYWLTMFRIRYHRR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGGSMŞYPLQLFCADGEYNSMAAAFFNTPEK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF PVVEHADDTQPARLQGLILRSQLIVLLKHKVF VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR
						VFKLFRALGLRHLVVVDNRNQVVGLVTRKD LARYRLGKRGLEELSLAQTGPKAQATAEGRV AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ QEEQGREVACGALQKNQEDSSRRVDLEVAR M
854	2204	A	7037	139	2604	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG VTQRDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDQDPDPQPRY TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFII DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLG\LP WALIFFSFASGTFQLVVLYLFSITSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	Α	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ .
862	2212	Α	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF YKNIQKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ CIKPNYQSPPKECDYNILANSVA
864	2214	Α	7214	845	1619	SDKGGKKADRKNHIRHAFPLLPHRVRERLH DPKVPVDADHVQGQDPGRAAHDIIIGEDVTE KVSKDPLAPDEVGDTDEGHDRHGHREVGQR HGHDQEEVAYEERACEGGKFATVEVTDKPV DEALREAMPKVAKYAGGTNDKGIGMGMTV PISFAVFPNEDGSLQKKLKVWFRIPNOFQSDP PAPSDKSVKIEEREGITVYSMQFGGYAKEAD YVAQATRLRAALEGTATYRGDIYFCTGYDPP MKPYGRRNEIWLLKT
865	2215	A	<i>7</i> 246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS LANMAKPRLY
866	2216	A	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL DLKKSDFSTRWQKQRCPVVKSKCRENASPFF FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCGPCPKNWICYKNNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYICM ORTV
867	2217	Α	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI MPFFQTLWLMNANRFCSIFTTTNVANNCWW TPYHCWLSVVVCRCESHGI
868	2218	A	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	Α	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHI.LA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF
870	2220	Α	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA
871	2221	Α	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR ALRGAALPGESEAGDPESLRSSVNADWIQYS

SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of peptide	hod	in NO:	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	(09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	[914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	·	Ì	ŀ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ł	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Į	ļ	Į.		peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
<u> </u>	<u> </u>		ļ	sequence		DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
ļ		j	}		j	PFIC
872	2222	A	7413	1061	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC
					 -	PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
		İ				CGLIKPCPANYFAFKICSGAANVVGPTMCFED
1		{				RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ
		1				KAFDMYSGDVMHLVKFLKEIPGGALVLVAS
		Ì				YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
İ	i i	ł	•			SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE GWPELLEMEGCMPPKPF
873	2223	Α	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ
""		١.,	/ 122		2331	DHPGQHCETPSLLKIERKLF
874	2224	A	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISTTE
			ļ			WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
		1				LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL
ļ	j ,	j	j .			GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
						AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
			•			MSSLNLDHWLKGAKREEWEPPPQSPALTHSP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
		[•			AEMOTPRRLGWGWYHTLTLYLWEEK
875	2225	A	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG
		}	, .	•		SERHEP*HGGVLFRLGPSAPPGKL
876	2226	Α	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
						ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	Α	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC
		ľ				TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGREILELVQKGIPV
				!		TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII
		İ				SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
		ļ				GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
						KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL
		1				DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
'				!		NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
878	2228		7586	315	1232	KGDAGENTALLEAGRSDSRHGGPIS ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
0/0	2220	Α	/360	313	1232	RGRMOAACWYVLFLLOPTVYLVTCANLTNG
						GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
		ľ				QTFRGKENDTDLDLRYDTPEPYSEQDLWDW
						LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
]						GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF
		1				RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
						AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
]					•	YKLVQKVCPDYNYHSDTPYFPSG
879	2229	Α	7605	479	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
880	2230	A	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG
	,					GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
1		1				VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
						MAGNTISIFPTMMVCMMAWRPIQALMAISAT
						FKMLESSSQKFLQGLVYLIGNLMGLALAVYK
						CQSMGLLPTHASDWLAFIEPPERMEFSGGGL
881	2231	A	7615	291	1452	LL SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
901	1022	Λ.	(013	£31	1432	NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
[SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL
}						LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
						CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI
						QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL
L		L		L		TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted begimning .nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIFTICFVPYHAFRFIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST
882	2232	A	7617	67	379	SEFKPGYSLHDTSVAVKIQSSSKST RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
884	2234	A	7638	2640	2861	AVATVGPISVAVGASHVFFQFYKKGKHLSS APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP
885	2235	A	7642	201	455	GLSCHTSHSG PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A .	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707		2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lvsine, L=Leucine,
seq-	иепсе	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			***	amino acid	of-peptide	T=Threonine, V=Valine, W=Tryptophan,
1	<u> </u>]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ	1	peptide	sequence	/=possible nucleotide deletion, \=possible
1						nucleotide insertion
	ļ		<u> </u>	sequence		EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
i	ĺ	ĺ	j			i e
i	1	ļ				GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
		ļ				DESTKTKDQILTSRINAVERDLLEPSPADQLG
1	ţ	ĺ	(NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV
l	}					TREPARRLFLFGEEPSKLDQDVLAALECADV
1	ł					DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
} ·		l	ľ			KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
L						GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240	Α	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
891	2241	A	7721	61	1175	KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL
1		ł				VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
1	1					HSPSDWITSHPEAPODFEOFFSDPYRKTPSPN
						KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
		1				YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
		1				AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
		}				WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
						KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR
						SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
						EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
						LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
ļ		j				EAFKEWKEWIIKCLAVLQK
892	2242	Α	7723	2	1650	
092	2242	Α .	1123	2	1030	SAPTAPARPCRAERGSGGGMLALLAASVALA
ļ		}				VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
1						PAMPMQGGAQSPEEELRAAVLQLRETVVQQ
						KETLASARAIRELTGKLARCEGLAGGKARGA
						GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
						DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ
1						LRETVVQQKETLASARAIRELTGKLARCEGL
						AGGKARGAGATGKDTMGDLPRDPGHVVEQ
[]		ĺ				LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
						FREVLQQRLGELERQLLRKGAELEDEKSLLH
1						NETSAHRQKTESTLNALLQRVTELERGNSAF
1						KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
						ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
						WGNNPIELLINDKVAQLPLFVSDGKWHHICV
1						TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
1						KPGGVLILGQEQDTVGGRFDATQAFVGELSQ
{					1	FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
						NNVDVFGGASKWPVETCEERLLDL
893	2243	Α	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
[DNYNDTSLVENHLCPATEGPLMASFKAVFVP
						VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET
						FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
('			LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
j :						HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
						LFAKVSQGHHNNSLPRCTFSQENQAETHAWF
			}	i		, ,
į l						TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
						QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
J J			j			IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL
				į		GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
						CTGPASLCQLFPSWRRSSLSESENATSLTTF
894	2244	Α	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
					ĺ	VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
						SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
j						VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
<u> </u>			[D
895	2245	Α	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
L						LWLSLFLHAGKEAPHCPRTRPL
896	2246	A	7754	1	372	SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVSTAALAVILCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775 '	85	496	SCOTTOPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	Α	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPIFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP ÄQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796		807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				residue of peptide	sequence	/=possible nucleotide deletion, \=possible
ļ		ļ	ļ	sequence	1	nucleotide insertion
903	2253	A	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
***		' '	'**	-	50.	VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
ļ		l	.	ļ		RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
		1				ARLLYESRKRGMLENCILLSLFAKEHLQHMT
		1	•			EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
}		ļ]]]	EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
						LFEKPR .
904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
		[AGARLTGWTMNVFRILGDLSHLLAMILLLGK
						IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
		ł	1	}		ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
					ŀ	WTFSIYLESVAILPOLFMISKTGEAETITTHYL
		ŀ				FFLGLYRALYLANWIRRYQTENFYDQIAVVS
1				}	}	GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
ļ ·	·		i I			RSYSSI
905	2255	Α	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
						QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
					j	QVEQNLELMTKRAVKAENHVVKLKQEISLL
						QAQVSNFQRENEALRCGQGASLTVVKQNAD
						VALQNLRVVMNSAQASIEQLVSGAETLNLVA
006	2056		7000			EILKSIDRISEVKDEEEDS
906	2256	Α	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
						HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
						PAAPAAPPFSSSGVLGTGLCELDRLLQELNA
						TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
ł i			1			RPSLPSSPSPGLPKASATSATLELDRLMASLSD
						FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
						KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
			l			NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
						GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
						RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
						FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE
						GRPLCENHFHARRGSLCATCGLPVTGRCVSA
						LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
						COPCFLKLFG
907	2257	A	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
L						YCKSQAWG
908	2258	Α	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
				,		SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET
						YTCPKMIEMEQAEAQLAELDLLASMFPGENE
						LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
						NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
					·	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
						SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
						SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
}						KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
				,		EEKVFSVNGARGNHMDFGQLYQFLNTKGCG
						DVFQMFLWV
909	2259	Α	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
l						LISSEILLIPSKYLFESK
910	2260	Α	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
[[[PSHRVNAEPGCVVTNACASGPCPPHANCRDL
						WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
ļ İ						SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
	i				LI	QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK

SEQ ID NO: of nucl- eotide seq- uence	NO: of peptide seq-uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLELNKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAGPGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERRGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRCRIGIKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEFKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQLVPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYALLLQHLPVLWWLPRYPVRDWLLGDLLSGLSVA IMQLPGCLSARAFATHQ WRTWLQCSRARAYALLLQHLPVLWWLPRYPVRDWLLGDLLSGLSVA IMQLPGCLSARAFATHQ WRTWLQCSRARAYALLLQHLPVLWWLPRYPVRDWLLGDLLSGLSVA IMQLPGGLSSRSSRSMAFDILGI SFGFTASVFMI PFHG
911	2261	Α	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL
912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PWPQGAVHFNLLDVAKDWNDNPKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL
						VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHKWIIAPKGFMANYCHGE CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD ECGCG
913	2263	A	7892	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	A	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN RLSNITRPFFSKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	Α	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK INCSWFIRANPGEIITISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCI.P ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	ł	09/496	соптекропой	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	1		peptide		/=possible nucleotide deletion, \=possible
_				sequence		nucleotide insertion .
]]	<u> </u>]	VILRETDEKLDGTGYGDYVKIYDGLEENPHK
1		•		İ		LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV
l		1		i .		NAARGFNATYQVDGFCLPWEIPCGGNWGCY TEOORCDGYWHCPNGRDETNCTMCOKEEFP
		1		[CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
l			l	ļ		COPGNFHCKNNRCVFESWVCDSQDDCGDGS
	ł					DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
			l			CTCKLYSLRMFERRSFETQLSRVEAELLRREA
		l				PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
	İ	ŀ		ľ		RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
	<u> </u>	ļ	}			RSRHSGSLALVSADGDEVVPSQSTSREPERNH
İ		1	1			THRSLFSVESDDTDTENERRDMAGASGGVAA
1	1	İ	1			PLPQKVPPTTAVEATVGACASSSTQSTRGGH
	1					ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG
			l	•		REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
						ASDQGQGLRQPYNATNPGVRPSNRDGPCERC
	İ	}				GIVHTAOIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
1		l	ł	Ì		GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQQTEDVLAA
						GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
						VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE
		1		İ		HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
ļ	ļ	l	ł	}		RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS FSSSWPSAQHLTPSVFNPW
921	2271	Α	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF
'2'	22/1	^	1757	012	012	MMVYSFRALSFKESTWATFOHGGEATKSRSL
	<u> </u>					SSTQ
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
ŀ	ļ					VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN
						WELVKI'N
923	2273	Α	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
		ĺ	ļ	ļ		LLPLLLLPAGCRALEETLMDTKWVTSELAWT
i	ľ	ĺ	l			SHPESGWEEVSGYDEAMNPIRTYQVCNVRES
ļ	į					SQNNWLRTGFIWRRDVQRVYVELKFTVRDC NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
						MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
J		J	}			FGPLSKAGFYLAFODOGACMSLISVRAFYKK
]		1				CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
1		l	1	, , , , , , , , , , , , , , , , , , ,		AVEVSVPLKLYCNGDGEWMVPVGACTCATG
		j]			HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
] .	l	Ì			NSRTTSPAASICTCHNNFYRADSDSADSACTT
			ļ			VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
·						LLYNVICKKCHGAGGASACSRCDDNVEFVPR
						QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
1	•					GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
		1	j		ļ ,	IASTVTSQMNSVQLDGLRPDARYVVQVRART
		l				VAGYGQYSRPAEFETTSERGSGAQQLQEQLP
						LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS
						DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA
}			}			VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
						KQPGRREVFVAIKTLKVGYTERQRRDFLSEA
		1	1			SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
]			NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
		[[KYLSEMNYVHRDLAARNILVNSNI.VCKVSDF
]			GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY
L	L	L	L	L		ATMATISASDY WSTOTYM WEYMSTOERPT

SEQ ID NO: of NO: of NO: of nucl- ectide seq- uence uence	·
eotide sequence ue	·
sequence under the last amino	
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=Serine amino acid of peptide T=Threonine, V=Valine, W=Tryptop	
amino acid of peptide T=Threonine, V=Valine, W=Tryptop	
peptide /=possible nucleotide deletion, \=poss	
sequence nucleotide insertion	
WDMSNQDVINAVEQDYRI.PPPM	
LMLDCWVRDRNLRPKFSQIVNTI	
SLKVIASAQSGMSQPLLDRTVPD'	
WLDAIKMGRYKESFVSAGFASFI	
EDLLRIGVTLAGHQKKILSSIQDM	IRLQMNQT
924 2274 A 7985 1 503 FRPRTKKATAMYLEHYLDSIENL	DCTL ODNE
924 2274 A 7985 I 503 FRPRTKKATAMYLEHYLDSIENL QLMRELDQRTEDKKAEIDILAAE	
PDORVERLOKIONAYSKCKEYSD	
QTYEMVDKHIRRLDADLARFEAI	
SDFESSGGRGLKKGRGQKEKRGS	
EEDTPKKKKHKGG	
925 2275 A 7994 447 589 LPCSFCAQCMSSFERVWLQQSHF	HNPRWNSR
SPIRCYCQHWPHCVHC	
926 2276 A 7996 925 582 GPCKVCCTTLAIMLQCHSFYRKD	VQVEHPKS
LNPKYSQIENFLSADMALKRKCL	LSISDLDFW
IWDAQPVGIMQTLQNLKKIPNPG	CFWSQAFQI
RDTQPILPLGGRYYITIRQ	
927 2277 A 7998 2 353 RIQRPLNSRSPNHSLFVKAELTAK	
CLLLVTLALCCYQANAEFCPALV	
SEPLFKLSLAKFDAPPEAVAAKLO	JVKRCIDQ
MSLQKRSLIAEVLVKILKKCSV 928 2278 A 8004 130 588 LAPLRCQPGTRTQPRSHPAANDP	SAAMSAAG
2278 A 8004 130 388 LAPLR-CQF-GTR-TQF-RSHF-AARDF-A	
GPRTVFFWAPIMKWGLVCAGLA	
LSTAQSAVLMATGFIWSRYSLVII	
VNFFVGAAGASQLFRIWRYNQEI	
929 2279 A 8007 2 1016 EFARRVFIAAREMSLLRSLRVFI	VARTGSYP
AGSLLRQSPQPRHTFYAGPRLSAS	SASSKELLM
KLRRKTGYSFVNCKKALETCGGI	
HKEAQKEGWSKAAKLQGRKTKE	
GNITVLVEVNCETDFVSRNLKFQ	
GTMMHCQTLKDQPSAYSKGFLN	
GPDREGSLKDQLALAIGKLGENN	
KVPSGFYVGSYVHGAMQSPSLHI ALVICETSEQKTNLEDVGRRLGQ	
SVGSLDDEPGGEAETKMLSQPYL	
YVQPQGVSVVDFVRFECGEGEEA	
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARK	QNRNSKEL
GLVPLTDDTSHAGPPGPGRALLE	
PGGRRKDWSCSLLVASLAGAFO	SSSFLYGYN
LSVVNAPTPYIKAFYNESWERRH	
TLLWSVTVSIFAIGGLVGTLIVKM	
HTLLANNGFAISAALLMACSLQA	
GRFIMGIDGGVALSVLPMYLSEIS	
QVTAIFICIGVFTGQLLGLPELLGK	
GVIVVPAVVQLLSLPFLPDSPRYL	
RAVKAFQTFLGKADVSQEVEEVI IRLVSVLELLRAPYVRWQVVTVI	
CGLNAIWFYTNSIFGKAGIPPAKII	
IETLAAVFSGLVIEHLGRRPLLIGG	
GTLTITLTLQDHAPWVPYLSIVGI	
PGGIPFILTGEFFQQSQRPAAFIIAC	
FAVGLLFPFIQKSLDTYCFLVFAT	
YFVLPETKNRTYAEISQAFSKRNK	
DSAVTDGKINGRP	
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFG	
NRNARRKAAPRIECSHIRHAWDH	
AEMGLAVDPNRAVPLRKRKVKA	MEVDIEER

NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glu	
nucl- peptide in nucleotide location F=Phenylalanine, G=Gly	ntamic Acid, vcine, H=Histidine,
eotide seq- USSN location corresponding l=lsoleucine, K=Lysine,	
seq- uence 09/496 correspondi to last amino M=Methionine, N=Aspa	
uence 914 ng to first acid residue Q=Glutamine, R=Argini	
amino acid of peptide T=Threonine, V=Valine,	
residue of sequence Y=Tyrosine, X=Unknow	vn, *=Stop codon,
peptide /=possible nucleotide de	letion, \=possible
sequence nucleotide insertion	
	AEASLPEKKGNTLSRD
	DYKAMARDEKNYYQD
	PAEWQDFLDSLQKRK
932 2282 A 8011 412 1 SNLCLGNSWRWRWA	KSRHHCIPTVTLSKRSG
	RVPGKETARVLRAGK
	PPPPPGSPGPGCRQFHQ
	IRGKVKMRRALRRAPA
STRASSROPNPK	NOX TEMEGRAPI
	KDLAGRLPAGPRGMGT
	VRESVFTVEGGHRAIFF
	LHFRIPWFQYPIIYDIRA
RPRKISSPTGSKDLQM	IVNISLRVLSRPNAQEL
PSMYQRLGLDYEERV	/LPSIVNEVLKSVVAKF
NASQLITQRAQVSLLI	RRELTERAKDFSLILDD
	EAKQVAQQEAQRAQF
	AEGEAEAAKMLGEAL
	QNISKTIATSQNRIYLTA
934 2284 A 8023 255 982 SOFSLSOVLVDSAEEC	SSLAAAAELAAQKREQ
	RKLNHQEVVEEDKRL
	WELKEEEKKKECAARG
	AERWERKKKRKNPDLG
	TKQIKPDMETYERLRE
	THVPSTEEIDRMVIDLE
KQIEKRDKYSRRRPY	NDDADIDYINERNAKF
NKKAERFYGKYTAEI	
	NSLAWTVTSYVFLKFL
	LTLLGIGPSHRHSLSIRL
SQHSSPAPMYSQTFHI	
	FKLLLIGDSGVGKTCVL
	IDFKIRTIELDGKRIKLQ YYRGAMGIMLVYDIT
	EHASADVEKMILGNKC
	KLALDYGIKFMETSAK
	KAKMDKKLEGNSPQG
SNQGVKITPDQQKRSS	
937 2287 A 8039 393 311 EETIHSENSYILEKYIP	ISANLTLTIA
938 2288 A 8052 675 1334 LHPAATSTAWLHVPP	GLSMALSWVLTVLSLL
	PITNATLDRITGKWFYI
	IQATFFYFTPNKTEDTIF
	TYLNVQRENGTISRYV
	KTYMLAFDVNDEKNW
	LGEFYEALDCLRIPKSD
939 2289 A 8055 12 1039 SSVAEFPERVOLSOPO	NWNFSGAGGAWSLDF
1	ESIMDGKQGGMDGSKP
	MGDAVSDWSPMHEAA
	/AVNIITADHVSPLHEA
	GAQVNGVTADWHTPL
	LQHGASVQPESDLASP
. , , , , , , , , , , , , , , , , , , ,	LIAYGGNIDHKISHLGT
	KLLESGADVNQGKGQ
	ACLLMDFGADTQAKN
	AQLFLEREGPPSLMQL
	JTKLVLPEDLKQFLLH
L L COSO A COSO	000000000000000000000000000000000000000
	SEPSQPSQPSQPGGHLI
AKLKIMDLHLFDYSE	PGNFSDISWPCNSSDCI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	İ	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ł	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
]	j					ANSVVVWVNIQAKTTGYDTHCYILNLAIADL
						WVVLTIPVWVVSLVQHNQWPMGELTCKVTH
\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \						LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS
			1			RKKMVRRVVCILVWLLAFCVSLPDTYYLKT
ľ			1			VTSASNNETYCRSFYPEHSIKEWLIGMELVSV
		ľ				VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS
						RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI PFTCRLEHALFTALHVTQCLSLVHCCVNPVL
ł				1	}	YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA
						SRVSETEYSALEQSTK
941	2291	Α	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS
						PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV
						IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ
						KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	Α	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS
						VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLIATANQGSEATYESGF
						VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA
						GROGTYEQYFGPGTRLTVTEDLKNVFPPEVA
						VFEPSEAEISHTQKATLVCLATGFYPDHVELS
						WWVNGKEVHSGVSTDPQPLKEQPALNDSRY
						CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE
'						NDEWTQDRAKPVTQIVSAEAWGRADCGFTS
						ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	Α	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI
'	2275	••	5070	•	0,,,	TASERLRRRPRATARLRAHAAPPEPPLAVFAP
						PSDRKELLALPVACDPVIASVMSWVQAASLI
						QGPGDKGDVFDEEADESLLAQREWQSNMQR
						RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
						EVILNYGRLRGTLSALLSWCHLHNNNSTLINK
						INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA
!						EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK
						PKPHMDFGTDSQF
944	2294	Α	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK
						MAATSGTDEPVSGELVSVAHALSLPAESYGN
						DPDIEMAWAMRAMQHAEVYYKLISSVDPQF
<u> </u>						LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK
					į	SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA
						VYISVQDKEGEKGVNNGGEKRADSGEEENT
, l						KNGGEKGADSGEEKEEGINREDKTDKGGEK
						GKEADKEINKSGEKAM
945	2295	Α	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL
[İ					SADRRVLGLREWGRPASERECSLCQRLKREL
						NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT
						GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW
ļ				l	1	GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE
946	2296		8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF
	~	• •	0001			VAIFA VPLILGO EYEDEERLGEDEYYOVYY
	ļ			ļ	l	YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
						DITEAIETTISLETARADHPKPVTVKPVTTEPQ
[[ļ					SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC
L			0001			KKVGRRLLMTLWMGVWQEEIGR
947	2297	Α	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF
L				l		SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- ucncc	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
948	2298	B	8093	3905	846	PCPPAIMYQSSNKC MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLIKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPTMAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A		9	2374	ARRADTVLLEŠPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD NPMSTPKDSNPKA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRILLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133		1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFTYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALVLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	Α	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide scquence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
			·			DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	Α	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS QAGSLV
958	2308	Α	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP. LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960	2310	A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN AGANLQNYGETSPDAISTNSEGAQENHDDLM SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST EFLTSSLMNIQHFLEDETVATVMPMKIQVSNTKINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
961	2311	A	8172	1442	682	SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN FPEFSFDFTREQLMEENESLKQELAKAKMAL AEAHLEKDALLHHIKKMTVE TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI
701	2311	^	01/2	1442	002	ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG EVQVSDKERHTQLEQMFRDIATIVADKCVNP ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS VLRRMQKKYWKTKQVFIKATGKKEDEHLVA SDAELDAKLEVFHSVQETCTELLKIIEKYQLR LNGMKS
963	2313	Ą	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGFGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV AWLSRAEWDQVTVYLFCDDHKLQRYALNRI TVWRSRSGNELPLAVASTADLIRCKLLDVTG GLGTDELRLLYGMALVRFVNLISERKTKFAK VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	Α	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLQELSAEERQKYLE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQQPKKK RSFSLSFSLLSPSEMMALGAAGATRVFVAMV AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ RCYCGEGLSCRIOKDHHOASNSSRLHTCORH
966	2316	A	8207	416	4082	RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPQFFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHILREDDIHIYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQVWFVSWRESFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATIHENIG AAGFKMSP
969	2319	A	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR VPGRLLLLLPALCCLPGAARAAAAAAGGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALRWEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR SASRSRSTRSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321		8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRRREAKAPR MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFFSTCGAGLQTILTGAPRLL QAIARDGIVPFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI LERKAKSRQVGKEKGKYKEELIEKMQE
973	2323		8237	279	4610	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEIINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHIPPRLPAAIFR PPPGSGLGNLGAGGTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
974	2324	Α	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL
						G

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion.
975	2325	A	8249	sequence 62	1571	nucleotide insertion I.VALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT IILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA
976	2326	A	8257	298	7086	GNMACWPOLRLLLWKNLTFRRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTFMENSQEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEBSDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD SMKVAVDGLALNFYEGGITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTILSTHHMDEADVL GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEIFLKVAEESGVDA ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

NO. of mucl- colide seq- uence	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence coide sequence per per per per per per per per per pe						B .	
Costion 1944 1944 1945						T .	
uence 91496 1914	1	1			L.		
1914 ng no first amino acid residue of peptide peptide peptide peptide peptide sequence Particular of peptide peptide sequence Particular of peptide peptide sequence Particular of peptide peptide sequence Particular of Particula		_	1	l .			
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence per sequence			j				
residue of peptide pep			ł		, –		
peptide sequence Possible nucleotide deletion, impossible							
sequence	1	•	l		peptide	•	
PAYFVCIAL/YSLIVPPEGYYSLELQPWMY NEOTYPYSMDAPEDITELLIALITEDEGG TRCMEGNPIPDITCO, AGEEWTTAPYPQTIM DLFONGWITMONPSPACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKKMLPV CPPOAGGILPPGRACQCSSINKKMLPV CPPOAGGILPPGRACQCSSINKKMLPV CPPOAGGILPPGRACQCSSINKWLPN NTQALPPSGEVNDATK, GWKKHLKLAKDSSA DEFLINKLGENATIOLOTIONNYKWENNIGW HAISSELAVINNALIRANIQKGEPSSHYGITAF NIPILIALITKOQLSEVAPATTSVDVIV, VSICVIPA MSFVPASFVVPLIQERVSKAKHLQFISGVKPVI YVISNIPVAMILITKOQLSEVAPATTSVDVIV, VSICVIPA MSFVPASFVVPLIQERVSKAKHLQFISGVKPVI YVISNIPVAMILILLILYGWSTPLMYPASFVFK NESTAVVDISVALLEGGIGSVATPTLELFOR KLINNIDLIKSVFLIPPHCI, GRGLIDMYKNO AMADALEEFGENEVEYSLEVENDLORGUNLFAM AVEGVVFFLIVLIQYRETRERPVAKLSISHLA AVEGVVFFLIVLIQYRETRERPVAKLSISHLA BEDEDSVERERQRILDGGONDILEIKELTITY RREKKPAYDRICVGIFPGECFGLLGVNGLAK SSTFKMLTGDITYVIGADAFLNRNSLISHLEV HQNMMYCCPGDAFTELLTGREHVEFFALLRG VPEREVGKVGWARRLOFWVFILEETTIM SGGRKRKLSTAMALIGGGONDILEIKELTITY RREKKPAYDRICVGIFPGECFGLLGVNGLAK SSTFKMLTGDITYVIGADAFLNRNSLISHLEV BERKEKRSTAMALIGVGFOLAFFGSVVF EKHNNMLQYQLISSLSSLARFISLSQKKRLH IEDYSVSQTTLDQVFVNFAKLOGDDDHILKDL SHNOTYVOTAVALTSTQDEKVKERSV TLLRGVSIIGTIGAGIFSRKGVLQNTGSVGM SLITMYTVOTAVALTSTQDEKVKERSV TLLRGVSIIGTIGAGIFSRKGVLQNTGSVGM SLITMYTVOTAVLTSTQDEKVKERV TLLRGVSIIGTIGAGIFSRKGVLQNTGSVGM SLITMYTVOTAVLTSTQDEKVKERV TLLRGVSIIGTIGAGIFSRKGVLQNTGSVGM SLITMYTVCGVLSIFGALSYAELGTITKKSGGH YTYTLEVFGPLFAFVRVWVELLIRPAATAVIS LAFGRYLLEFFTCCEPLELAKLTTAVGHTVW VINSMSVSVSARQIGITFCKLTALILIIVPOV MCILKGQTONFKDAFSKGROSSITRL-LAFYY MYAYAGWFYLNFVTEEVENPEKTIPLACISM AVITIGYVTLINVAYFITHAEELLISNAVAVT FSERLIGRIFSLAVFFFVALSCFGSSMNGGYFAV SREPLYSARSGHLFFTGTCSVGVVPPLALAGGSGSE EAAAGTEGDAQGWPGAGSSADQDDEEGVVK FORSLWPWDSVRNILBSALTEMCVLYDVLSI VPOKKFMTLDFVSQDALPFKNOFQCHILQLIG GTETTYPPLSMSOGNAHRRALAGRIDFSGGSGSE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FORSLWPWDSVRNILBSALTEMCVLYDVLSI VPOKKFMTLDFVSQDALPFKNOFQCHSKLLQ RDFMSELLRLRQHWKLRKVGDKILGDLSTYNS AGSLFAVGHKKRATAKCHERSPUCHEVHALL RDFMSELLRRGHWKLRKVGDKILGDLSTYNS AGSL			ļ	1	, · ·		nucleotide insertion
NEQTTFVSNDAPEDTGTLELLNALTKDPGFG TRCMGRPIPDITCO,GEEWTIAPYOTIM DLFONGNWTMONPSPACOCSSINKIKMLPV CPGAGGLPPONKONTADILODIORINISDY LWKTYVQILAKSIKNIWVNEFRYGGPSLGVS NTQALPPSGEVNDATIO,GOMENISTIKAKOSSA DEFLINSLGEFMTGLDTINNNYWWFNINKOW HAISELAVINAULA RAND, (KGEPPSFIYOTIAF MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVAPMTTSVDVLVSICVIFA MEPLINLTKOQLSEVARMICSVAFTHVIFICORSEVAF MEPLINLTKOQLSEVARMICSVAFTHVIFASFVVK WVLSINVALLLLLYGWSITPLMYPASFVK MEPLINLTKOQLSEVARAMICSVAFTHVILLYGYRFIRERRYNAKLSSLVN AMADALERFGENRYVSLSEVALIVATION AMADALERFGENRYVSLSEVALIVATION AMADALERFGENRYVSLSEVALIVATION AMADALERFGENRYVSLSEVALIVATION AMADALERFGENRYVSLSEVALIVATION WYLSKOKKVOEWAMRKLOLKVYGEKYAONY SGGKKRKLSTAMALIGGPPVYFLDEPTTOMD PKARRELWKCLSVYKGERSVVLTISHMEC BALCTRIMAMVAGRERCLGSVOHLKNIRFGD GYTIVLYNAGSNPOLKYGERSVVLTISHMEC BALCTRIMAMVAGRERCLGSVOHLKNIRFGD GYTIVLYNAGSNPOLKYGERSVVLTISHMEC BALCTRIMAMVAGRERCLGSVOHLKNIRFGD GYTIVLYNAGSNPOLKYGERSVVLTISHMEC BALCTRIMAMVAGRERCLGSVOHLKNIRFGD GYTVLEVYOLFONFAKOOSOBOHLKDL BENSTSSELCFIPPEVYMPKERVYSTISKGG VYLESVGTTLDQVVVALUSERSMESV BENSTSSELCFIPPEVYMPKERVYSTISKGG VYLESVGTTLDQVVALUSERSMESV SERLINGVALUSERSSELLSTATIONS SVAN SLINGVANGKLISPSCOHLONGVAVLEN SLINGVANGKLIGENSTATIOLILEVYPEDIAL VANKOR PREPICE PLAIKLITAVGTTVM VLNMSVSWSNARIOIETTICGIPTHTLTOWAYLFILACISM AND ALVERVYSNIKALISPSCOHLONGLIGGPTTL VDKKPRWERMSEKITRTLOILLEVYPEDIAL VDKKPRWERMSEKITRTLOILLEVYPEDIAL VDKKPRWERMSEKITRTLOILLEVYPEDIAL VDKKPRWERMSEKITRTLOILLEVYPEDIAL VDKKPRWERMSEKITRTSLIELSCOMSCHIPTIOLILSK KRIGGIGENT SAN TENGOR SORD BERNSTS SCHOOL STAN TEN					·		LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
TRÉMEGNPIPDITEO, AGEEWTTAPYPOTIM DLFONGWITMONPS ACQCSSISKIKKMI PV CPPOAGGI.PPPORKQNTADIL QDLTGRNISDY LYKTYVQIJAKSIKMIKWVERPYGOPSLOYS NTQALPPSOEVNDATKOMKKHIKLAKDSSA DREINSLGEPATIG DITENNYXWYERPYGOPSLOYS NTQALPPSOEVNDATKOMKKHIKLAKDSSA DREINSLGEPATIG DITENNYXWYERPYSOPSLOYS HASSELWVDNALIKANIQKGENPSHYCITAF NEPIJILTKOQLSEVARMITISVDVILVSICVIFA MSFVPASFVYRLIGURGVSKAKHLQIESGVKPVI YWLSNFFWDMONTVVPATLUIIPICTQOKSY VSSTNLPVALLLLLYGWSITPLMYPASFVFK IPSTAYVVLTSVALFIGINGSVATFVLELFTDIN KLNINNDILKSVPLIPPICLGRGLIDMYKNQ AMADALERFGENRFVSPLSWDLVGRNLFAN AVEGUVYFLITVLLQYRFIEPRRVNAKLSILN DEDEDVRRERGRILDGGQNDILEIKELTITY RRKRKPAVDRICVEIPGECFGLLGVNAGK SSTFKMI.TGDITVIRGDAFLNRSILSNIHEW HONMGYCPGDAITELLTGREHVFFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPVVPLDEFTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRAMMVNGRFRCLGSVOPHLKNFGD GYTIVVRIAGSNPDLKPVQDPFGLAFFGSVPK EKHRMIMOQOLPSSLSARFISLSGSKKRLH IEDYSVSQTTLDOYFVNRAKDSDDDHLKDL SJERNGTVVDVAVALTSFI ODEKVKESVY 1		ļ		ļ			
CPPGAGGLPPOKKONTADILODLTGRNISUS LVKTYVOJIASLAKNIKIWNEPETYGGPSLOVS NTQALPPSQEVNDATKQMKKHLKLAKDSSA DERLINGGREWIGLOTRINNEV VISICUFA NTQALPPSQEVNDATKQMKKHLKLAKDSSA DERLINGGREWIGLOTRINNEV VISICUFA NTQALPSQUEVALLAKINGGEPPSHYGITAF NIPLINITKQQLSEVAPMTISVOVJSCIVFA MSFVPASFVVPILOERVSKAKHLQFISOVRPVI YWLSSFVWDMCNYVVPATLVIIFICPQOKYV YSTRILVLALLLILUTGWSTFINITYPASFVFK IPSTAYVLTSVNLFIGINGSVATFVLELFINI KLININDLIKSVFLIPPHCLGRGKDLIDMVKNQ AMADALERGENRIVSPLSWDLVGRNLFAM AVEGVVFFLITVLIQVSFFREPPVAKLSPLN DEDEDVRRERQRILDGGGQNDILEIKELTKYI RRKRFAVDRICCIGPPGECFGLLGVKGAGK SSTFKMLTGDITVTIKGDAFLNRNSILSNIHEV HQNMGVCPPDATITELLTGREHVEFFALLRG VPEKEVGKVGEWARRLGLVKYGEKYAGNY SGGNKRLSTAMALIGGPVVPLGFFGALBFGSVFK KEHRMALGVQLVSSLSSLARIESIS,GSKKRLH HEDVSVSQTTLDQVFVNFAKDQSDDDHLKNFG GYTIVVRIAGSNPDLKFVQDFFGLAFFGSVF KEHRMALGVQLVSSLSSLARIESIS,GSKKRLH HEDVSVSQTTLDQVFVNFAKDQSDDDHLKNFG GYTIVVRIAGSNPDLKFVQDFFGLAFFGSVF KEHRMALGVQLVSSLSSLARIESIS,GSKKRLH HEDVSVSQTTLDQVFVNFAKDQSDDDHLKNFG GYTIVVRIAGSNPDLKFVQDFFGLAFFGSVF KEHRMALGVQLVSSLSSLARIESIS,GSKKRLH HEDVSVSQTTLDQVFVNFAKDQSDDDHLKNF KEHRMALGVQLVSTLSSLDGEKVREFYV STRIMLGVQLVFFGLAFFGSVF KEHRMALGVQLVFFRAKDGSDDHLKDV SGRYKLSTAMALIGFPVVLTSHQEKVRESVV 1124 TARGVSLEFFFQCFVTAVRKFPVVSTISKGG YLQGNVXNGLFSCLGSKVSSVS SLTIWTVCGVLSFACVALSTISKGG YTLQGNVXNGLFSCLGSKVSSTTRLPLAFYV MQLKGGTQNFKDAFSGROSSTTRLPLAFYV MQLKGGTGNFKDAFSGROSSTTRLPLAFY MAYAGWFYLTNAYFTTINAEELLLSNAAVV VLNSMSVSWSARGIFLTECKLTALLIIVFOV MQLKGGTGNFKDAFSGROSSTTRLPLAFSFTC LFMVALSLYSDPFVALSCFGSMAGGVFAV SRLFYVASREGHLFFILSIMHVRKTFPLPAUSH AVTIGVYLTNAYFTTINAEELLLSNAAVV SRLFYVASREGHLFFILSIMHVRKTFPLPAUSH AVAGUFYLTNAYFTTINAEELLLSNAAVV SRLFYVASREGHLFFILSIMHVRKTFPLPAUSH AVAGUFYLTNASFTGGSVTUTLFALISM AVTIGVYLTNAYFTTINAEELLSNAAVV SRLFYVASREGHLFFILSIMHVRKTFPLPAUSH AVAGUFYLNASKSGSSADQDDEGVVK PROSLUKGRAFMANTAGARDFFQCSGGEVE EBAAGTEGDAAGWFQAGSSADQDDEGVVK PROSLUKGRAFMANTAGARDFFQCSGGEVE EBAAGTEGDAAGWFQAGSSADQDDEGVVLP LADAASMSGVVAVRISIESACKEVPFEGLL VRDKKFMILDPVSQOALPFKQNFLCKEI FAQLSREAVQKSQVFHIVVKNQIJGGFFSUL LFKRPLPSKKFGSFHWOTTLEAAGKKEPDVCPL DVQPSBLEGGSAYKNSIQKQFAFEQCCFCEPHLTVULE EFAQLSREAVQKSQVF	1	1		1			TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
LVKTYVOJIAKŠLKNŠKIWVNEĖRYGGPSILOS NTOAI PPSQEVINDATKQMKKHKLAKDSSA	İ		ĺ				DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
NTQAIPPŠQEVNDATKQMKKHLKLAKDSSA DREINSIGRPMTGLOTRINNVEWPNINGW HAISSFLINVINNAILRANLQKGENPSHYGITAF NIPILNLTKQQLSEVAPMTISVDVLVSICVIPA MSFVPASFVVFILQERVSKAKHLQFISGVKPVI YWLSNFVVDMCNYVVPATLVIIIPICCQQKSV VSITILLYGNSTIPLAHIPICATION KINNINDILKSVILEPPIECLGRGILDMVKNQ AMADALERGENRVSPISWDLVGRNLFAM AVEGVVFFLITVLQYNFFIRPPPIACLGRGILDMVKNQ AMADALERGENRVSPISWDLVGRNLFAM AVEGVVFFLITVLQYNFFIRPPPIACLGRGILDMVKNQ AMADALERGENRVSPISWDLVGRNLFAM AVEGVVFFLITVLQYNFFIRPPPIACLGRGILDMVKNQ AMADALERGENRVSPISWDLVGRNLFAM AVEGVVFFLITVLQYNFFIRPPPIACLGRUMGAGK SSTFKMLTGDITVTIRGDAFLNRAISILSNIHEV HQNMOYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGGKYAGNY SGGNKRKLSTAMALIGGPPVVFLDETTIGMD FKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRIAMINVNGFFRCLGSVQHLKNREFOUND FKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRIAMINVNGFFRCLGSVQHLKNREFOUND GYTIVVRIAGSNPDLKPVQDFFGLAPGSVPK EKHRMILQYQLPSSLSSLARIFSILSQSKKRIL EDVSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV TLRGVSIIGTTIGAGFISPKGVLQNTGSVGM SLHKNQTVVDVAVLTSFLQDEKVKESYV TLLRGVSIIGTTIGAGFISPKGVLQNTGSVGM SLHKNQTVVDVAVLTSFLQDEKVKESYV TLLRGVSIIGTTIGAGFISPKGVLQNTGSVGM SLTIVTVCGVLSLGALSTAELGTTIKKSGGH YTYLEVFGPLPAFVRVWVELLIRRAATAVI VLNSMSVSWSARIQIFLTFCKLTALLIIIVPGV MQLKGQTQNFKDAFSGROSSITRLPAFYYQ MQLKGQTQNFKDAFSGROSSITRLPAFYYQ MYAYAGWFYLNFVTEWEPREKTIPLACISM AVGITYVATNVAYTTINAEELIANAVAVT FSERLLGNRSLAVPIFVALSCFGSMNGGYPAFVYTFI LPMVALSLYSDPFSTGIGVTILTGVPAYYLFI WDKKPRWFRIMSEKTTRTQIILEVYPEGOKL VAGLIVLRYKCPDMFPRFVKUPELAFFTC LPMVALSLYSDPFSTGIGVTILTGVPAYYLFI WDKKPRWFRIMSEKTTRTQIILEVYPEGOKL VAGLIVLRYKCPDWFARSPROSTALFSFTC LPMVALSLYSDPFSTGIGVTILTGVPAYYLFI WDKKPRWFRIMSEKTRTTQIILEVYPEGOKL VAGLIVLRYKCPDWFARSPROSTALFSFTC LPMVALSLYSDPFSTGIGVTILTGVPAYYLFI WDKKPRWFRIMSEKTRTQIILEVYPEGOKL VAGLIVLRYKCPDOADLPPKONPTLALISTIC LPMVALSLYSDPFSTGIGVTILTGVPAYYLFI WDKKPRWFRIMSEKTRTQIILEVYPEGOKL VAGLIVLRYKCPDOADLPPKONPTLACIST AGSISTANAVGRIGGAPPIGDLGTVN LPKRPLPSKRYCPSPHOWTLLCEG FAQLSREAVQIKSQOPHIVVKNQISQFFTES AGSISFPHOTFEVIKNTDLDLDKREP							CPPGAGGLPPPQRKQNTADILQDLTGRNISDY
DRFLINSLGREMTGLDTENNINK/WEININGW HAISSFLINVINALIRANILOGENSHYOTION MIPPLALIT QOLSEVAPMTTSVDVLVSICVIPA MSPVPASFVVLIGENVSKAHLOFISGVKPW YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY VSSTRILPVLALLLLLYGWSITPLMYPASFVFK IPSTATYVLTSVAHLFGINGSV ATFVLELFIDN KLINNINDLKSVFLIFPHEGLGRGLIDMVKNQ AMADALERGENFVSHLEWIOLGRNLEAG AVEGVVFFLITVLIQYRFIRPRPVNAKLSPIN DEDEDVRRERGRILDGGGOMDLIEIKELTKIY RRRKRPAVDRICVGIPPGECGGLIGVNGAGK SSIFKMLTGDITVIRGDAFLOKGSLSPIN DEDEDVRRERGRILDGGGOMDLIEIKELTKIY RRRKRFAVDRICVGIPPGECGGLIGVNGAGK SSIFKMLTGDITVIRGDAFLITGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKGEKYAGNY SGGINKRLSTAMALIGGPPVFLDEPTITOMD PKARRFLWKALSVVKEGRSVLTSHSMEEC EALCTRMAMVRIGRFRCLGSVQHLKNRFOD GYTIVVRIAGSNPDLKPVVGEKYAGNY SGGINKRLSTAMALIGGPPVFLDEPTICMD PKARRFLWKALSVVKEGRSVVLTSHSMEEC EALCTRMAMVRIGRFRCLGSVQHLKNRFOD GYTIVVRIAGSNPDLKPVVGEKYAGNY SGGINKRLSTAMALIGGPPVFLDEPTICMD PKARRFLWKALSVVKEGRSVVLTSHSMEEC EALCTRMAMVRIGRFRCLGSVQHLKNRV SGGINKRLSTAMALIGGPPVFLDEPTICMD PKARRFLWKALSVVKEGRSVVLTSHSMEED FKHRINLQYVDAVLTSHLQDEKVKESYV TLRGVSIIGTIIGAGFISFKGVLQNTGSVGM SLTIWTVGGVLSJCLGKLSVKGENGVF TLLRGVSIIGTIIGAGFISFKGVLQNTGSVGM SLTIWTVGGVLSJFGALSVSGTIKKSGGG YTVLLRGVSIIGTIIGAGFISFKGVLQNTGSVGM SLTIWTVGGVLSJFGALSVSGTIKKSGGG YTVLLRGVSIIGTIIGAGFISFKGVLQNTGSVGM SLTIWTVGGVLSJFGALSVSGTIKKSGGGVF TVLLRGVSIIGTIIGAGFISFKGVLQNTGSVGM VLNSMSVSWSARIGITTCKLTALIIIVPGM MQLIGQTONFKDAFSGRSGSTEFICLLSMHYKKHTTLPATVY LIFHTMIMLFSGDLDSLLNTETLATYVG MYAYAGWFYLNFVTEEVENPEKTIPLACISM AVITGSVLTINVAYFTITIAGTITVM VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIST VGNVSSTIPVASREGHLPBLSMHYKKHTTLAVIT LIFHTMIMITSGDLDSLATINGT TRATTITUT GTOTT VGNTAGT REGGILSTAMALAQRIDPSQGSGSEE EAAGTEGDAQRWPAGSSSADDDEGGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDULSI VRDKKFMTLDPSQDALFPROPTOLILIK KKSLLASNDAVILCKEI FAQLSREAGRAVGIKSQVPHVVKNQIIGGFTNN LFKRFLKSKFSSPHWOTKLEAGANVLLCKEI FAQLSREA	1	i	ĺ	ĺ			LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS
HAISSEINVINNALIRANILOKGENPSHYGITAF NIPPIALITA (DILGEVAPAMTEVUVSICIVIPA MSPVPASFVVPILIQERVSKAKHLOPISGVKPVI YWLSNFVWDMCNYVVPATLVIIIPICPQOKSY VSSTNLFVLALLLLLYGWSTIPLMYPASFVKF IPSTAYVVLISVNLFIGINGSVATFVLELFIDN KLINNIDLIKSVELIPIPHCIGGRGIJDMVKNQ AMADALERGENRFVSPLSWDLVGRSILFAM AVEGOVPFILIVILQYRFIFIRPYNAKLSPLM DEDEDVRRERQRILDGGGONDILEIKEIKTY RRKRKPAVDRICVGIPPGEGEGELIGVNGAGK SSTIFKMLTGDITVIRGDAFLNRNSILSNIHEV HQMMCYCPOPDATTELLTGREVFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKVAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMVNGRRFCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVFK EKHRNMLQYQLFSSLSSLARIFSILSQSKKRLH IEDVSSQSTLDQVFVNAKDDDDHLKDL SLHKNGTVVDVAVLTSFLQDEKVKESVY 1000 1000 1000 1000 1000 1000 1000 1							NTQALPPSQEVNDATKQMKKHLKLAKDSSA
NHPLNITKQQLSEVAPMTISVDVLVSICVIPA MSPVPASPVVJQENVSKAKHLQFISGVKPV YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY				}			DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
MSFVPASFVVELIQENVSKAHLQFISGVKPVI YWLSNFVWDMCNYVVPATLVIIIFICPQOKSY YSSTNLPVLALLILLY GWSITPLMYPASFVFK IPSTAYVVLTSVNLFIGINGSVATFVLELFITDN KLNNINDILKSVFLIPPHICLGRGLIDMVKNQ AMADALERGGENFFVSPLSWDLVGRILFAM AVEGOVFFLITVILVIQFRFIFREFVNAKLSPLN DEDEDVRRERGRILDGGGQNDILEIKELTKY RRKRKPAVDRICVGIPPGECFGLIGVNGAGK SSTIFKMLTGDITVIRGDAFLNKNISISNIHEV HQMMCYCPGPDAITELLTGREFVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMMYGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKFVQDFFGLAFPGSVFK EKHRNMLQYQLFSLSSLARIFSILSQSKKRLH IEDVSNSQTILDQVFNVRAKDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 1000 1000 1000 1000 1000 1000 1000 10		1	İ				HAISSFLNVINNAILRANLQKGENPSHYGITAF
YWLSNFWDMCNYVYPATLVIIIPCPQOKSY VSSTNLPVLALLILLY QWALLILLY VSSTNLPVLALLILLY QWALLILY RSTALYVLTSVNLFIGINGSVATFVLELFIDN KLNNINDILKSVFLIPPHECLGRCLIDMYKNQ AMADALERFGENRY VSPSLWDLVGRNLFAM AVEGVVPFLITVLJQYRFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGGONDILEIKELTKIY RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSIFKMLTGDITVIRGDAFLNKNSILSNIHEV HQNMGYCPGPDATTELLTGREHVEFFALLRG VPEKEVGKVGEWARRLU KYGEKVAGNY SGGNKRKLSTAMALIGGPPVVFLDEPITGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMDVNGFRELGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFGSVPK EKHRNMLQYQLPSSLSSLARFSILSQSKKRLH IEDYSVSQTTLDQVFVNFADQSSDDHLKDL SLHKNGTVDVDAVLTSFLQDEKVKESYV 100 1567 PGSTTSFSLCFIPPCVFTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TILLRGVSIIGTIGAGIPSKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLEVFGDPLPAFVRWVELLIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIGHETLSHALTAVGITVVM VLNSMSVSWSARIGHETLSHALTAVGITVM VLNSMSVSWSARIGHETLSHALTAVGITVM VLNSMSVSWSARIGHETLSHALTAVGITVM VLNSMSVSWSARIGHETLSHALTAVGITVM VLNSMSVSWSARIGHETLSHALTAVGITVM VLNSMSVSWSARIGHTLSHALTAPYYG MQLIKGGTQNFKDAFSGRDSSTTRLJAFYYG MQLKKGGTQNFKDAFSGRDSSTRLPLAFYYG MQLKKGGTQNFKDAFSGRDSSTRLPLAFYYG MQLKKGGTGNFKDAFSGRDSSTRLPLAFYYG MQLKKGGTGNFKDAFSGRDSSTRLPLAFYYG MYATAGWFVLNFVTEEVENPEKITPLACISM AIVTIGYVLTNVAYFITTNAEELLLSNAVAVT FSERLLGNFSLAVPFVALSCFGSMNGGVFAV SRLFYVASREGGHLPEILSMHVRKHTPLPAVIV LIFHLTMIMLFSGDLDSLLNFLSFRAKHFGLA VAGLYLLRYKCPDMHRPKYVLFFRALFSFTC LFMVALSLYSDFFSTGGFVITLTGVAYYVLFI WDKKFRWFRMSEKITRTGUILEVYPEEDKL PADAASMSGVRAVNISISACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSIK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDEPNSELLARLRQHWKLRKYGDKLIGDLSYVS AGSLFPHNGTFEVIKNTDLDLDKKIPEDYCPL DVQRPSDLELGSAYIKVSIQKQAPDIGDGTVN LFKRPLPKSKPGSPHYMYKLLCKEI FAQLSREAVQIKSQVPHIVVKNOILSGPFSLQ LSISLCHISKDKKSQKFATENQCPEDHLYVLE LISLCHISKNFKSSPHYMYKNLEAAAQNIVLLCKEI FAQLSREAVQIKSQVPHIVVKNOILSGPFSLQ LSISLCHISKNDKKSQKFATENQCPEDHLYVLE LISLCHISKNDKKSQKFATENQCPEDHLYVLE	1		ļ				NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
VSSTNLPVLALLLLLYGWSITPLMPASSEVEK BYSTAYVVLTYSVNLFGINGSVATTPVLELFIDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ AMADALERFGENRRVSPLSWDLVGRNLFAM AVEGVVFFLITVLJQYKFIRPRRVNAKLSPLN DEDEDVRRERQRILDGGGONDILEIELTKIY RRKRKPAVDRICVGIPPGEGFGLLGVNGAGK SSIFKMLTGDITVTIRGDAPLNRNSISNIHEV HQNMGYCPOFDATIELLTGRIVEFFALLRG VPEKEVGKVGWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLEDETTGMD PKARRLWNCALSVVKEGRSVVLTSHSMEEC EALCTTMAIMVNGFFRCLGSVOPLINRRFGID GYTTVVRIAGSNPDLKFPVOPFGLAPFGSVVK EKHRNMLQYQLPSSLSSLARIFSILSGSKRLH IEDYSSVGTTLDQVFVNFACDSODPHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCTIFPCVFTMVRKFVVSTISKGG YTUGNNYAGISPDLKFPVGPGEKVQLKKKV TLLRGVSIIIGTIIGAGIPISPKGVLQNTGSVGM SLTIMTVCGVIJSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRWVELLIIRPAATAVIS LAFGRYLLEFFTQCEIPELKLITAVGTVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLJAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAVFTITNAEELLLSNAVATI FESERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGFILPBILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNSLSCFGSMNGGVFAV SRLFYVASREGFILPBILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNSLFSFR WLFIGLA VAGLIYLRYKCPDMHRFFK VPLFIPALFSFTC LFWVALSLYSDFFSTGJGFVITLTGVAYYYLFTL WKKFRYLFRSKFGSSRNGGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARALERGENVILGULSVRS AGSLFPHGTFEVIKNTDLDLKKREDEVVLE ROFNSELLARLRQHWKLRKYGDKILGDLSYRS AGSLFPHGTFEVIKNTDLDLDKKREDOYCH, DVQNPSDLEGSAYIKVSIQKAQPDIGDLGTVN LFKRPLPKSKFGSPHQVTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQILSGPFSLQ LSISLCHISNBVKSQKATERQCFPEDHLYVLE	<u> </u>	ĺ					MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
PSTAYVVLTSYNLFIGINGSVATFVLELFTDN KLNNINDILKSVFLIFPHFCLGRGIDMVKNQ AMADALERFGENRYSPLSWDLVGRNLFAM AVEGVVFFLITVLJQYKFFIRPRFVNAKLSPLN DEDEDVRRERQRILDGGGONDILEIKELTKIY RRKRKPAVDRICVGIPFGECFGLLGVNGAGK SSTFKMLTGDITVIRGDALNKNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVOGWAIRKLGLVKYGGEKYAGNY SGGNKRKLSTAMALIGGPVVYFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTEMAIMVNGFRECLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAPFGSVFK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTILDQVFVVRAKDQSDDBHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV PGSTIFSELGFIPPFCVFFIWRKPVVSTISKGG YLQGNVNGRLPSLGNEEPFGQEKVQLKRKV TILIRGVSIIGGTIGAGGIFFKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYTLEVFGPLPAFVRWWELLIRPATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTALIIVPGV MQLKGGTORPFROAFSGRSTRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGVVLTNVAYFITINAEELLLSNAVAVT PSERLLGHFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLFEILSMHVRKFIPLAVIST LAFGRYLLEPFIQCEIPELAIKLITAVGTTVVM VAGLIYLRYKCPDMHRPFKVPLFFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPLAFSFTC LFMYALSLYSDPFSTGIGFFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PALASANSGVRAVRISIESACERQVFFFGGLL VAGLIYLRYKCPDMHRPFKVPLFTFCSWFFFGGL FFQLSSRLSAGKPPLRISFFGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKCPDHRPFKVPLTSFFGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKCPDMHRPKVPLTSFFGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKCPDMHRPKVPLTSFFGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKCPDHRPKPVLTSFFGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKGPDAPLRSFFTGSWGVLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKGPGDAPLRSFFTGGSWGLPP LADAASNSGVRAVRISIESACERQVFFFGGLA VAGLIYLRYKGPGDAPLRSFFTGGSWGLPP LADAASNSGVRAVRISIESACERQVFRFGGLA VAGLIYLRYKGPGSAGSADQDDEGGVVK FQPSLWWDSVRNNLRSALTERQCFFRFGGLI VRKKFMTLDPVSQDALPRGNFFTGGNKLQ RDFRYSELLRLRQHWKLRKYGDKILGDLSYRS AGSLFPHGTFEVIKNTDLDLDKREPDVCPL DVQFPSDLEGSAYIKVSGVGAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISGPFFSLQ LISSLCHISNDKKSQKFAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI FAQLSREAV		İ		ļ			YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY
KLNNINDILKSVFLIPPHECLGRGLIDMYKNQ AMADALERFGENRFVSPLSWDLVGRNLFAM AVEGVVFPLITVLIQYRFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGGONDILEIKELTKIY RRRKRFAVDRICVGIPPGEGFGLLGWNGAGK SSTFKMLTGDITVTIRGDAFLNRNSILSNIHEV HQNMGYCPGPDATTELLTGRIVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGMKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGFFRCLGSVOPLKNRFGD GYTIVVRILGSSPPDLKPVOPFGLAPPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSGKKRLH IEDYSSVGTTLDQVFVVNFAKDQSDDDHLKDL IEDYSSVGTTLDQVFVVNFAKDQSDDDHLKDL SLHKNQTVDDVAVLTSFLQDEKVKESYV 11LRGYSIIIGTILGGAIFISPKGVLQNTGSVGM SLTIWTVCGVISLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEPELALTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEPELALTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEPELALTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEPELALTTIKKSGGH YTYLLEVFGPLPAFFRGVENTATALIITVPGV MQLIKGQTQNFKDAFFSGRDSSITLLAFVYG MYAYAGWFYLNFVTEEVENPEKTIPLAFTYG MYAYAGWFYLNFVTEEVENPEKTIPLAFTYG MYAYAGWFYLNFVTEEVENPEKTIPLAFTYG MYAYAGWFYLNFVTEEVENPEKTIPLAFTYG MYAYAGWFYLNFVTEEVENPEKTIPLAFTYG LFLTMMLFSGDLDSLLNFAFRAFWLFGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGGFVITLTGVYAYYLFIT LFLTMMLFSGDLDSLLNFAFRAFWLFGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGGFVITLTGVYAYYLFIT LFLTMMLFSGDLDSLLNFAFRAFWLFGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGGFVITLTGVYAYYLFIT LFLTMMLFSGDLDSLLNGAFRFFKVPLFFBLESCHL VAGLIYLRYKCPDMHRPFKVPLFFDALFSFTC LFMVALSLYSDPFSTGGSGVTLTGVYAYYLFIT LFLTMMLFSGDLDSLLNGAFRFFKVPLFFDCHLQ GTETYLPPLSMSQNLAKLRQDFSGGSGSGVLP LADASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLAKLRQDFSGGSGSGE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQESLWPWDSVRNINSALTEMCVYDVLISK KKSLAGAQULLKGAERLTKSVTENGENLQ RDENSELLARCHWKLRKYGDKLIGDLSYRS AGSLPPHHGTFEVIKNTDLDLDKKREPDYCL DVQPSDLEGSAYIKVSIGKAQPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISGPFSSLG LISSLCHISKNOKKSQKAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQISGPFSSLG			ŀ	ĺ			VSSTNLPVLALLLLLYGWSITPLMYPASFVFK
AMADALERGERRFVSPLSWDLVGRNLFAM AVEGVVFFLITVLIQYRFIRPRPVNAKLSPLN DEDEDVRREQRILIQGGQNDILEIKELTKIY RRKRRPAVDRICVGIPPGECFGLLGVNGAGK SSTIFKMLTGDITVINGDAFLNRNSILSNIHEV HQNMGVCPQFDAITELLTGREHVEFFALLRG VPFEKEVGK VGEWARKLGU.VKYGER,VAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTISHSMEE EALCTRAMDMYNGRFCLGSVQHLKNRRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQVQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV P77 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVPTMVRRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLIRGVSIIGTIIGAGFISKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARQIFLTFCKLTALIIIVPGW MQLIKGGYQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAFTINAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPELSMHVKHTPLPAYVIL LHPLTMIMIFSGOLDSIVTLTGVPAYVIFI WDKKPRWFRIMSEKITRTLQIILEVPPEDKL LHPLTMIMIFSGOLDSIVTLTGVPAYVIFII WDKKPRWFRIMSEKITRTLQIILEVPPEDKL LHPLTMLAGRIFFSGOSGSEE EAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTIQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLIQ RDFNSELLRLRQHWKKFGOKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLERLRQHWKKFGOKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLERLRQHWKKFGOKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKKFGOKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRLRQHWKROGNLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCL DVQIPSBLELRRQHWKHONIGOPFPSLQ LISSLCHSSNDKKSGKFATEKQCYBEILGTVN LFKRPLPKSKPGSPHVVTKLEAAQNVLLCKEI FAQLSREAVQUKSQVPHIVVKNONISQPFPSLQ LISSLCHSSNDKKSGKFATEKQCYBEILGHTVNLE]	!	ļ				IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN
AVEGVVFFLITVLJOYRFFIRRPYNAKLSPLN DEDEDVRRERQRILDGGQNDILEIKELTKIY RRKRKPAVDRICVGIPGECFGLLGVNGAGK SSIFKMLTGDITVIRGDAFLNRNSILSNIHEV HQNMGVCPOFDATELTREEHVEFFALLGG VPEKEVGK VGEWARKLGL VKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRRGD GYTIVVKIAGSNPDLKPVQDFFGLAFPGSVPK EKHRMLQVQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESVV 777 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTIWNRFVVSTISKGG YLQGNVNGRLPSLGNKEPFGQEKVQLKRKV TLJRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLFFFIQCEIPELAIKLITAVGITVWM VLNSMSVSWARIQFITTCKLTALIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNNAYFITTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMGGGVFAV SRLFYVASRGGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDPFSTGIGVTILTGVPAYYLFII WDKKPRWFRIMSEKITRT,QIILEVVPEEDKL FMVALSLYSDPFSTGIGVTILTIGVPAYYLFII WDKKPRWFRIMSEKITRT,QIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLIG GTETT_LPPLSMSQDLAARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDEEGVVK FQPSLWPMDSVRNNLRSALTEMCVL,VDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLIKSVTENQENKLQ RDFNSELLRLQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPESDLEGSAYIKVSJQKQAPDIGGLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVYKNDISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1	İ			-		KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ
DEDEDVRERGRILDGGGONDILEIKELTKIY RRKRKPAVDRICVGIPPGEFGGLLGVNGAGK SSTFKMLTGDITVTRGDAFLNRNSILSNIHEV HQNMGYCPGFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWALSLSVKGERSVVLTISHSMEC EALCTRMADMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRNMLQVQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVDVAVLTSLDGEKVKESYV 1567 1567 1763TISFSLCFIPPCVPTMVRKPVVSTISKGG YLQGNVNGRLFSLGNKEPFQQEKVQLKRKV TLLRGVSIBIGTIDGAGFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTYLEVFGPLAFFNVWELLIRPAATAVIS LAFGRYLLEFFIQCEIPELAIKLITAVGITVVM VINSMSVSWSRQIFILTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AUTTIGVYLTNVAYFITIAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMMLFSGDLDSLNFLSFARWLFIGLA VAGLIYLRYKCPDHRIPPKVPLIFPALFSFTC LFMVALSLYSDPFSTGIPTITGVPAYVILFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL FAULSTSPLFSTGGSGSEE EEAAGTEGDAQGWPGAGSSADQDEEGVVK FQPSLWPWDSVRNINRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLDKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLDKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLDKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLDKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLDKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLELRRQHWKLRKVGOKILGDLSYRS AGSLFPHTGTFEVIKNTDLDLOKKIPEDYCL DVQPSGBLEGAAYIKSSQYPHTVYKNONISQFPFSLQ LISSLCHSSNDKKSGKFATEKQCYEDHLYVLE	1			ļ			AMADALERFGENRFVSPLSWDLVGRNLFAM
RRKRRPAVDRÜCVGIPPGEGELLGVNGAGK SSTFKMLTGDITVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLUKYGEKYAGNY SGGNKRLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGFFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRMLQYQLPSSLSSLAHFSILSQSKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVFTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TILRGVSIIIGTIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAKLITAVGITTIKKSGGH YLTMIVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAKLITAVGITTVVM VLNSMSVSWSARIGILTFCKLTALLIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVGEVENPEKTIPLAICISM AIVTIGYVLTNVAYFITINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLEISMHVKRHTTPLAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMRPFKVPLFPFALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PROKEN PROKENSER PROKENCYPFFEGLL LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWWDSVRNARSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPRQPNCTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLKKPEDYCPL DVIPPSDLEGSAYIKVSIGKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVUKSQVPHIVKKNIISQFPFSLQ LSISLCHSSNDKKSQKAPTEKQCPEDHLYVLE FAQLSREAVUKSQVPHIVKNINGIPSDYCPL	1	Ì	}	İ			AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
SSTEKMLTGDITVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGJ.VKYGEKYAGNY SGGNRKKLSTAMALIGGPPVFLDEFTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDVSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFFPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPFGQEKVQLKRKV TLLRGVSIIIGTIJGAGFPSFWGVLQMYGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYTLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEFFFIGCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFITINAELLLSNAVAVT FSEKLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRINGCGLLSSRLSAGKPPLRTSFTGSWGVLPP LADAASMSGVRAVBISESACFKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKLIGDLSYRS AGSLFPHIGTFEVIKNTDLDLKKIPEDYCPL DVQIPSDLEGSAYIKVSIGKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQKSQVPHIVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLTYVLE	1						DEDEDVRRERQRILDGGGQNDILEIKELTKIY
HONMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGERSVVLTSHSMEEC EALCTRMAIMVNGFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRMLQYQLPSSLSSLARIFSILSQSKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIEPPCVPTMYKRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFQCEIPELAIKLITAVGITVVM VINSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGVVLTNVAYTITNAEELLISNAVAYT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVKHTTPLFAVIV LHPLTMIMLFSGGLISLNIFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSSTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSSKITRTLQIILEVVPEEDKL PHWALSLYSDPSTTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSSKITRTLQIILEVVPEEDKL GTETYLPPLSNGNILARLAGRIDFSQGSGSE EEAAGTEGDAQEWPGAGSADODDEEGVVK FQPSLWPWDSVRNNLARLAGRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADODDEEGVVK KSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHIGITEVIKNTDLDLDKKIPEDYCPL DVQIPSDLCGSSAYKVSIOKQAPPIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVYPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						
VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPITIGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMYNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSRPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNJFAKDQSDDDHKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIIPPCVPTMYRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIJGGIFISFKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTVILLEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLKGGTQNFKDAFSGRDSSTIRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT SEERLLGNFSLAVPIFVALSCGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSQDLDSLLNFLSFARWLFGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGIVILTGVVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWQVLPP LADAASMSGVRAVRISIESACEKGVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPMDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNFQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAJIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWOTKLEAAQNVLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCFEDHLYVLE	1						
SGGNKRKLSTAMALIGGPPVYFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQVQLDSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNGTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVYTMYRKPVYSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISFKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSAGIPFLTFCKLTAILIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSPDFTGGFVTITTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFFEGLL RIRGGGLLSSRLSAGKPPLRTSFTGSWVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALALQRIDFSQGSGSEE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALALQRIDFSQGSGSEE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALALQRIDFSQGSGSEE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALALGKLIGDLSYRS AGSLPPHHGTFEVIKNTIDLDLDKKIEDDYCL DVGIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWOTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFFSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						
PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFADQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMYRKPVVSTISKGG YLQGNYNGRLPSLGNKEPPGGEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YYYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLEPFFIQCEIPELAIKLITAYGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITR.PLAFYYG MYAYAGWFVLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRINGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASNSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPRKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWGTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKKSQKFATEKQCPEDHLYVLE	ļ		ŀ				
### FALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLFIPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTTCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFLNFVTEEVEPHPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPBILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIVLRYKCPDMHRPFK VPLFIPALFSFTC LFMYALSLYSDPFSTGIGFVITLTGVPAYVLFI WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLCFQSERCVFFFEGLI RHRGGGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRRKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGABRLIKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSVRS AGSLFPHRGTTEVIKNTDLDLDKKIPEDYCPL DVQRSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1	1					
GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVFTMVRKPVVSTISKGG YLQQNVNGRLPSLGNKEPPGGEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTVLEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTALILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTITNAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LIHLTMIMLFSGDLDSSLNFLSFARWLFIGLA VAGILYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRINGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGEE EBAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQFPSDLEGSAYIKVSIQQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKPATEKQCPEDHLYVLE	ĺ	İ					
SKHRNMLQYQLPSSLSSLÄRIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEK VKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIGTIGAGIFISPKGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLDFAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTALIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVTILTGVPAYYLFII WDKKPRWIMSEKITRT.QIILEVPEEDKL VMCKPRWIMSEKITRT.QIILEVPEEDKL RRGGSLRCVLGKLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFRTILDPVSQDALPFKQNPGTUQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDVCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCFEDHLTYLE							
977 2327 A 8260 3 1567 IPGSTISPSLCFIPPCVPTMVRKPVVSTYV FOR SLIKKNQTVVDVAVLTSFLQDEKVKESYV FOR SLIKKNQTVVDVAVLTSFLQDEKVKESYV FOR SLIKKNQTVVDVAVLTSFLQDEKVKESYV FOR SLIFEALSSLCFIPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLITIWTVGGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSMARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLXFYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRICE STORGSLRCVLGKLLGQLLCFGSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSGGSGSE EAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFRTIDPVSQDALPPKQNPQTIQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRCPHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYYLE							,
977 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVPTMVRFVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYULEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGYFAV SRLFYVASEGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII MYKKPRFFRIMSEKITRTLQIILEVVPEEDKL PLADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNPQTILQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVODKULGDLSYRS AGSLFPHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1			ļ			1
977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PLADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLFRSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	ł						l
YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYTLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLPPHIGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLESSNDKKSQKFATEKQCPEDHLYVLE	000	0005	<u> </u>	00.60		15/7	
TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVUSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTI.QIILEVVPEEDKL PHADAASMSGVRAVRISIESACEKQVHEVGLD GTETTYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSJ VRDKKFMTLDPVSQDALPPKQNFQTLQLISK KKSLAGAAQILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKRGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	911	2327	Α	8260	3	1567	
SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLISNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRINGEGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQVFATEKQCPEDHLYVLE	!						l ,
YTYILEVFGPLPAFVRVWELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFITINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFIMSEKITRTLQIILEVVPEEDKL PROSENEVLE SAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVYPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1	i .					I
LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTTCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFITINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVFFEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							{
VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRI.PLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPPQTIQLISK KKSLAGAAQIILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVYPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							l I
MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFITTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							1
MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	ļ						
AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE			· ·				1 7 7 7
FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFK VPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1				'	!	1
VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						
LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE						•	
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	j						
RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	978	2328	A	8261	2	2165	
LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						l i
EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	ľ	1			ì	1	
VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1					,	
RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						
AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1				į] ,	
LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1					İ	FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
	1						LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
					<u> </u>		

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- ucnce	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ
979	2329	A	8289	2	1053	WNKMEGRNFVYKMELLMSALSPCLL FVWNPRGGRKRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLFVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	Α	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	Α	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFIYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC
		ı	0004		12.10	VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY KAYLQDLVEGMDFQGPGES
984	2334	A	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYANPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	Α	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RIRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	Α	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIKKYDQMAIFHCLFWPSLTLL.GGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH QFKLCIMRRSKGRAEKS SSVVEFSALSVSMACLSPSQLQKFQQDGFLVLEGGELSAERCYMANORIGERYAFADVRILLORT
						EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS APGTSFLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906		MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRRWLGDPEHL
993	2343	A			2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE DMVTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVVAPPTLPGRDEP YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQITDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV NFILSTTI
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
995	2345	A	8390	194	3421	PSSWDYRACLS AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	•	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
1	•	1		sequence	! 	nucleotide insertion
		 				DFLSMKQSPALAPEERCRRAGSPKPVLRADD
						NNMGNGCSQKLATANLLRFLLLVLIPCICALV
						LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI
		l				QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT
)				TDASLPGDQSHRNTSACMNITHSQCQMLPYH
		!			<u> </u>	ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY OHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE
	•				i	AAKEGCESVLGMVNYSWPDFLRCSQFRNQT
	•	İ				ESSNVSRICFSPQQENGKQLLCGRGENFLCAS
1]				GICIPGKLQCNGYNDCDDWSDEAHCNCSENL
						FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC
1						DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD
						KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG
		1				DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM
	İ	l			į	NLPYNSTSYPNYFGHRTQKEASISWESSLFPA
1		[LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP
		l				CRALCEHSKERCESVLGIVGLQWPEDTDCSQ
			İ			FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ
	İ	ł				CVLASRRCDGQADCDDDSDEENCGCKERDL
1		İ				WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS
1						DSSDEWDCVTLSINVNSSSFLMVHRAATEHH
		ļ				VCADGWQEILSQLACKQMGLGEPSVTKLIQE
		1				QEKEPRWLTLHSNWESLNGTTLHELLVNGQS
			İ			CESRSKISLLCTKQDCGRRPAARMNKRILGGR
		İ				TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW
		1				VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
					İ	DISETGYVRPVCLPNPEQWLEPDTYCYITGW
		l				GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK
1		l				TITTRMICAGYESGTVDSCMGDSGGPLVCEK
						PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS
						YFVEWIKRQIYIQTFLLN
996	2346	A	8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSRS
		ļ				FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR
						NRGFYPWGQYNRGGYGNYRSNWQNYRQAY
]						SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS
						DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS
	'				İ	KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK
						ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST
		[SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
1						SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
1					[YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE
						KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
				:		SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
						DKAKGRKESEFDDEPKFMSKVIGANKNQEEE
1						KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
		ľ				AVOEKSSSPPPRKTSESRDKLGAKGDFPTGKS
						SFSITREAQVNVRMDSFDEDLARPSGLLAQER
						KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
						SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
1						KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
1						HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
			·		;	LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS
						SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL
L	ı ,		L—		i	COCCOUNTIAND LIEUTICA COLLOWN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			714	amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFOFRARGRGWG
					!	RGNYSGNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE DDESGTENREEKDNIQPTTE
997	2347	Α	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS KQENLESLEATGREEEGGQGPPVTTKGVLLA LLMAGLALQPGTALLCYSCKAQVSNEDCLQ VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL VSKKANQQLNFTEAKEACRLLGLSLAGKDQ VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE VFMETSTMSTETEPFVENKAAFKNEAAGFGG VPTALLVLALLFFGAAAGLGFCYVKRYVKAF PFTNKNQQKEMIETKVVKEEKANDSNPNEES KKTDKNPEESKSPSKTTMRCLEAEV
1000	2350	Α	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT AAIFTADGNMISASTLMDILLMNDFKLVINKI AYDVQCPKREKPSNEHTAEMEHMKSLVHRL FTILHLEESQKKREHHLLEKIDHLKEQLQPLE QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWVYSWDIMEPVTYFITFANSMVFF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ HFDVQQYNKLKEDLAKAKESLKQARHSLCL QMQVEELNEKN
1001	2351	Α	8410	1400		VGFWERPLRSSRWFRRSLRRWEMLARAARG TGALLLRGSLLASGRAPRRASSGLPRNTVVLF VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV LYLRIMDPYKASYGVEDPEYAVTQLAQTTM RSELGKLSLDKVFRERESLNASIVDAINQAAD CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE RKKRATVLESEGTRESAINVAEGKKQAQILAS EAEKAEQINQAAGEASAVLAKAKAKAEAIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA KDSNTILLPSNPGDVTSMVAQAMGVYGALT KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFI.SAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIHNKHAG
1003	2353	Α	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	[in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence	1	09/496		corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	neixce	l	914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ļ	į.	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		Ì		peptide	Soquence	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						SLDTENIDEILNNADVALVNFYADWCRFSQM
			Í			LHPIFEEASDVIKEEFPNENQVVFARVDCDQH
					}	SDIAQRYRISKYPTLKLFRNGMMMKREYRGO
						RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
					1	KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
			i			LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY
1						LGAMTNFDVTYNWIQDKCVPLVREITFENGE
		İ])	ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL
						ISEKGTINFLHADCDKFRHPLLHIQKTPADCP
1				,		VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
1			1			HSGKLHREFHHGPDPTDTAPGEQAQDVASSP
7004	0264		0.400	010	205	PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
				Ì	1	ACAAARSPADQDRFICIYPAYLNNKKTIAEGR
		· ·				RIPISKAVENPTATEIQDVCSAVGLNVFLEKN
]	KMYSREWNRDVQYRGRVRVQLKQEDGSLC
1						LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA DOSLQOGEGSKKGKGKKKK
1005	2355	Α	8453	90	530	OSHETKMOSGTHWRVLGLCLLSVGVWGOD
1003	2333	Λ	0733	90	330	GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE
						ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF
1						SELEQSGYYVCYPRGSKPEDANFYLYLRARG
1						NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW
						KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS
1						LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC
L			i			QLCIFN
1007	2357	Α	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL .
i i						SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT
1						VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM
						GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA
1 1						FWWHNKGLALIFCILQSLALTWYSLSFIPFAR
						DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP
1 1						PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS
						DPRWGCVGPSMPTSTCLPGAVEASTTKASLP
1009	2359	A	8465	134	954	KCPVDSSLPTPEACFL ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP
1007	2335	^	0403	134	734	NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH
1 1	'	•				LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
						FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL
j						TKLLVHSSLVGSILSALSALVGFIILSVKQATL
						NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD
						CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL
1						RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
						HDCGYEELLTS
1010	2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA
1 1		1				HRVALCHLAGCQEQAAWYHTLQILFFLVSAY
}						FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT
1 1		į				LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL
						SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	Α	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
		!				GTLETQFTCPFCNHEKSCDVKMDRARNTGVI
, ,		i		ļ		SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG
1		Ì		l		PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC
1010	22.60		0.00			RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM
						RMKYGGQEFWADLNAMNVYETTEFDQLRR
<u></u>						LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SVIRI.IEEANSRGI.KEVRFMMWNNHYILHNS FFREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKRKEYMNRKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI QYEEVSNTVSI
1013	2363	Α	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	A	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2366	Α	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN
1017	2367	Α	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLI.LVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV
1010	2500	1	0310	324	054	VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRKKGGPCCRQ RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536]	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	Ā	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAJVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT WIVEFFANWSNDCQSFAPIYADLSLKYNCTG LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS TPTTVSDGENKKDK
1025	2375	A	8546	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPA CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS ALPTKTVKPVENKDDDDSIADFLNSDEEEDR VSLQNLKNLGESATLRSLLLNPHLRQLMVNL DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES
1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK
1028	2378	Α	8569	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY
1029	2379	Α	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
1032	2382	A	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ĺ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	İ	1	!	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ		peptide		/=possible nucleotide deletion, \-possible
		<u> </u>	<u> </u>	sequence		nucleotide insertion
			i			QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
l		l	ŀ			VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
		1				TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
	l		İ			MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
ł						GERAFMNRTVGEIMWGYKDPLVNLINKYFP
ŀ			!			GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI
i	l		İ			SRIHLVDKWNGLSKVDFWHSDQCNMINGTS
						GQMWPPFMTPESSLEFYSPEACRSMKLMYKE
ļ		ĺ	!			SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
		l				CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL
		!				AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
1		ļ				KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA
	}	l	İ			ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
						LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
		<u> </u>				GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	A	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
						FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	Α	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ
1	['	PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL
ì		ŀ				VNMGDRTSMVQDPGSQAPTSWISESQVFQTT
1						EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
}]	J	ļ			GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
						F
1035	2385	Α	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
ł	ŀ	ŀ	!			SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
ĺ						SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI
						HVYKKNGVGKVGDQILLAIKGQKKKALIVG
1						HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
	L	<u> </u>				KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV
1						TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
Í	ļ	Ì	!			VAYRPISASVLSRPEASRTGEGSTVFNGAQNG
1						VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
ļ	}					VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
1000	2000		0616		0044	AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	Α	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
l		l				GMVAHINNSRLKAKGVGQHDNAQNFGNQSF
1		ļ				EELRAACLRKGELFEDPLFPAEPSSLGFKDLG
1						PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
	ļ					CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
!	}	J]			QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
		1				PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
		ĺ	!			GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
		1				PONLLRLLRKAVERSSLMGCSIEVTSDSELES
	· ·					MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
		İ				RVRNPWGRIEWNGAWSDSAREWEEVASDIQ
		[1			MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL
1					,	TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
		[RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
		1				VVCTCLVALMQKNWRHARQQGAQLQTIGFV
1		}				LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI
			[FTNSREVSSQLRLPPGEYIIIPSTFEPHRDADFL
		1				LRVFTEKHSESWELDEVNYAEQLQEEKVSED
		[! !			DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
]	MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
]	ļ	ļ				KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
		l				LNSYEMRLVIEKAGIKLNNKVMQVLVARYA
1		1				DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
L	L	L	l		<u> </u>	GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1038	2388	A	8621	3	1494	RSRMARAPLGVLLLGLLGRGVGKNEELRLY HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF GGIETLRVPSELVWLPEIVLENNIDGQFGVAY DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT VFLFLIAQKIPETSLSVPLLGRFLIFVMVATLI VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCIQP
1039	2389	Α	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98		ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	520	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP
1042	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1 "00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	F .	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-		1			M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	09/496	correspondi	to last amino	
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ĺ	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		}		peptide		/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
					T	TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT
			1			LDMITSTDHVLEQDFWICFTFYSVKERQI
1043	2393	A	8688	359	17	GLKTRAPATPTFQREVLGPAKQDMQRRCPRI
		' -			1	GLMTSLLKPIKRRWRDYKRWKSGGFTGESC
ł.		ļ	ļ		ĺ	HHADTLGDRGGLQGDHSELLQWQKRILRTE
j		}		İ		GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS
1044	2394	^	0/18	292	1490	1
[[ĺ			GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS
		·	1			YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL
ļ				ļ	ļ	VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL
l	1				}	NLALADLLFALTLPIWAASKVNGWIFGTFLC
]		İ '	1			KVVSLLKEVNFYSGILLLACISVDRYLAIVHA
1			1			TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR
J]	ł]	RTVYSSNVSPACYEDMGNNTANWRMLLRIL
1		i '	1			POSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK
		1		ļ		HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM
Ì	١.	1				RTQVIQETCERRNHIDRALDATEILGILHSCLN
Ì						PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS
Ì	ì	1				RPSFVGSSSGHTSTTL
1045	2205	_	8724	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ
1045	2395	Λ	8/24	254	3184	
		·	İ			DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY
ì	1	ł	}			YGEICDNACPCEEKDGILTVSCENRGIISLSEIS
!	 		Ī			PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL
ļ		1	į			HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL
j			ļ			ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF
						GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL
i			:			DLRGNRLKLLPYVGLLQHMDKVVELQLEEN
İ		ļ .				PWNCSCELISLKDWLDSISYSALVGDVVCETP
ľ		1				FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP
1		1				LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP
l	[KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA
	:		j			,
						YQTKSPVPLECPTACSCNLQISDLGLNVNCQE
1	ł . :					RKIESIAELQPKPYNPKKMYLTENYIAVVRRT
						DLLEATGLDLLHLGNNRISMIQDRAFGDLTN
1						LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY
]]			NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS
						GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS
					ļ	LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG
						VLVDEVICKAPKKFAETDMRSIKSELLCPDYS
İ						DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
[PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA
1					j	AGLFVLVMKRRKKNOSDHTSTNNSDVSSFN
1						MOYSVYGGGGGTGGHPHAHVHHRGPALPK
}						VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN
I				:	ļ	
l .						SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP
l [']						QQPQQQPPPQLQLQPGEEERRESHHLRSPAYS
]				ļ]	VSTIEPREDLLSPVQDADRFYRGILEPDKHCST
I						TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH
						QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE
1						YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	Α	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT
1						AMAGALVRKAADYVRSKDFRDYLMSTHFW
(GPVANWGLPIAAINDMKKSPEIJSGRMTFALC
ļ						CYSLTFMRFAYKVQPRNWLLFACHATNEVA
						•
1015	0202	<u> </u>	0741	672		QLIQGGRLIKHEMTKTASA
1047	2397	Α	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP
				· ·		AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK
<u></u> _						PPTTKLLHSSPLWNFFAQQL
1048	2398	Α	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	!			peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						VAVPNGQPPSAARYMPREVPPRFRCQQDHK
1						VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN
						PNNAQVTGALLQSESGTAPDSTLGGAAASNY
	j					ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS
						DMEEWPCIASKDTESSSENTTDNNSASNPGSE
[KSTLPGSTTSNKGKGSQCQSASSGNECNLGV
Ì						WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS
						RKGALETDNSNSSAQVSTVGQTSREQQSKME
1						NAGVNFVVSGREQAQIHNTDGPKNGNTNSL
						NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ
1						STGDRKTGSVGSWGAARGPSGTDTVSGQSNS
						GNNGNNGKEREDSWKGASVOKSTGSKNDS
						WDNNNRSTGGSWNFGPQDSNDNKWGEGNK
						MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN
						SSTGAWDNQKGHPLLENQGNAQAPCWGRSS
ľ						SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY
1 .						RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG
						QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES
						AATQTKNSGGWGDAPSQSNQMKSGWGELS
						ASTEWKDPKNTGGWNDYKNNNSSNWGGGR
						PDEKTPSSWNENPSKDQGWGGGRQPNQGWS
						SGKNGWGEEVDQTKNSNWESSASKPVSGWG
						EGGQNEIGTWGNGGNASLASKGGWEDCKRS
1						PAWNETGRQPNSWNKQHQQQQPPQQPPPPQ
						PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA
1						TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP
						MTSKSASDSKSMQDGWGESDGPVTGARHPS
						WEEEEDGGVWNTTGSQGSASSHNSASWGQG
						GKKQMKCSLKGGNNDSWMNPLAKQFSNMG
						LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA
1 1					į	MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS
1						GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN
1						SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP
1 1					j	GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL
1						LLQQQQQQLLQNQRKISQAVRQQQEQQLA
						RMVSALQQQQQQQQRQPGMKHSPSHPVGPK
1 1					}	PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF
						SSGGMDYGMVGGKEAGTESRFKQWTSMME
			1		ł	GLPSVATQEANMHKNGAIVAPGKTRGGSPY
					İ	NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK
						IGSKSSNASWPPEFQPGVPWKGIQNIDPESDP
]						YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS
		ļ		į		NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK
				- 1		FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS
						RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG
				ł		WGTQDSRLASASTWSDGGSVRPSYWLVLHN
						LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF
						ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS
	j		j .			LETGQNQSDPVGPALNLFGGSTGLGQWSSSA
	·]			l		GGSSGADLAGASLWGPPNYSSSLWGVPTVED
						PHRMGSPAPLLPGDLLGGGSDSI
1049	2399		8748	200	1387	VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF
'''			3,		130,	LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS
						PODCGFKDHQPLTLQALTVELARWTLMLLLS
				ļ		TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT
}			1	ļ		ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA
	ı		ļ	ļ		PFALSALLYGANNNLVIYLQRYMDPSTYQVL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL MAAGACYAAGGLQVPGNTLPSPPPAAAASP MPLHITPLGLLLLILYCLISGLSSVYTELLMKR QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP GLLEGFSGWAALVVLSQALNGLLMSAVMKH
						GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA FFLATLLIGLAMRLYYGSR
1050	2400	Α	8758	3	1660	WVSSMGFEELLEQVGGFGPFQLRNVALLALP RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS HQDVWLEAHLPREPDGTLSSCLRFAYPQALP NTTLGEERQSRGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFFA GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN
1051	2401	A	8759	515	1625	EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARJIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH
1052	2402	A	8763	1106	70	RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR
1053	2403	A	8768	2	712	RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM
1054	2404	Α	8769	344	527	REATTLACRNSCWVFSRCSLGACKPTVCSMP SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QQESPAAGAARMNCKEGTDSSCGCRGNDEK
						KMLKCVVVGDGAVGKTCLLMSYANDAFPEE YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL ECSALTQKGLKAVFDEAILTIFHPKKKKKKRCS EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	А	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH RRDQKWHDKQYKKAHLGTALKANPFGGAS HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK NGKKITAFVPNDGCLNFIEENDEVLVAGFGR KGHAVGDIPGVRFKVVKVANVSLLALYKGK KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME TQSEPSELELDDVVITNPHIEAILENEDWIEDA SGLMSHCIAILKICHTLTEKLVAMTMGSGAK MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL DPKLLDARTTALLLSVSHLVLVTRNACHLTG GLDWIDQSLSAAEEHLEVLREAALASEPDKG LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTVHNPGLALLHLLLLYGLVVSTALI WHPINKLAALLLLPYLAWLTVTSALTYHLWR DSLCPVHQPQPTEKSD
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATIYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824	1	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNITSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV
1063	2413	A.	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR
1064	2414	A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
L				sequence		nucleotide insertion
i		i	i	•		VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
						SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
		1]			EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
[•		ļ	QAQEELAEMKRYSESSSKLEEDKDKKINEMS
		1	[KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
		ĺ				LQQQVKQLQNQLAECKKQHQEVISVYRMHL
1						LYAVQGQMDEDVQKVLKQILTMCKNQSQK
10/5	2415		0041			K
1065	2415	A -	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
1 .						APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL
1		ł	}			KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
ļ						TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
1			i			DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
!						DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
1000	2416		0052	2006	2204	AVLFITGIILTSGKCRQLSRLCRNHCR
1066	2416	Α	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG RRRRGRVVSRKKMSLKSERRGIHVDOSDLL
Į		Į į				j
1			i		į	CKKGCGYYGNPAWQGFCSKCWREEYHKAR
						QKQIQEDWELAERLQREEEEAFASSQSSQGA QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
						, -
1						VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE
1 .						EQSECAQDFYHNVAERMQTRGKVPPERVEKI
						MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
ľ						QKRIRALRWYTPQMLCVPVNEDIPEVSDMVV
1						KAITDIEMDSKRVPRDKLACITKCSKHIFNAI
						KITKNEPASADDFLPTLIYIVLKGNPPRLOSNI
1						QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
						KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
1						WSPDACLGVKOMYKNLDLLSOLNERQERIM
1 :						NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
1						KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	A	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1007	2417	^	0000	1372	1515	LRQAWATKQDPISKKK
1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	A	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	Ā	8866	293	1675	PYPOGGYPOGPYPOGGYPOGP
**/*	4720	**	5500	4.7.3	10,5	YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
						EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
						LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
						WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
						VALSVLTASLSYMVGMIASFYNTEAVIMAVG
						ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
						VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
						VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI
1				!		FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
						WHGSASCTSPLSCPQAQPREKDASLQPSCMY
						TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
]						HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
						EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
1						GDMRSGGLIPVLSPE
1071	2421	A	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
''''	4741	^	3000	~	330	DDKMGSNTFFKRNDCRYVMISCKADMAYDN
						VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII
						LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
1 10/2	2722	^	0070	ور	920	GKQHLLKTEKSKLLSDISARLWFTYRKFSPI
1 1						GGTGPSSDAGWGCMLRCGQMMLAQALICRH
						LGRDWSWEKQKEQPKEYQRILQCFLDRKDC
L		L				LOIW HO HERVREY! RE I VRILVOPEDRADO

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
İ				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł			!	peptide		/=possible nucleotide deletion, \=possible
<u> </u>	<u> </u>		ļ	sequence	<u> </u>	nucleotide insertion
		ļ	1		i	CYSIHQMAQMGVGEGKSIGEWVLGPNTVAQ
j		}				GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA RFPKKLCRVLPL\SADTAGLTGP
1073	2423	A	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
1073	272	1 ^	0073	140	712	*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL
		İ				RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
		**		,	.55	KEISFGDYICHTFQGDCWADRSPLHEAAAHG
						RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
			}			*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	Α	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
	İ					SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
1		ļ	ŀ			PWPSLLDKEREESLRQKRLSERERIGELGAPE
İ						VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
ŀ		ł				TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
						EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
						QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
			Ì			QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
[•					PRRGEIGLTR*RNCHHLNAQVM**VVSRIIRR
						MEAVRTAKREPESTVLMRREPLHPFNPRRET KERE
1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
1070	2420	, n	0022	140	169	*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
		İ				FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
						VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR
						VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
						LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
						RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
						QYPALHRAGTEWQLSALHRAPRSTQPDKAC
						RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
1050						\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	Α	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
						PALPFAATPGSRGQALCRGGRRRQHLHGPLH
1079	2429	Α	8912	121	376 ·	RP*QAAPALHAGCQLAPHPPT
10/9	2429	A	0912	121	3/0 .	NLIWKLCVTERRLVILDNYDLASE/YEANKYI
						CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY
1080	2430	Α	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
	50	I	5,20		-/	GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
						DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
]						YAPICMEYGRVTLPCRRLCQRAYSECSKLME
				l	ì	MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
				ļ	Į	GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
				ļ		SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
						IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
			İ			WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA
				ļ	-	STVTQGSHNKACTMLFMLYFFTMAGSVWW
		' !		l		VILTITWFLAAVPKWGSEAIEKKALLFHASA
						WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
				j	j	VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
						VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
1001	2731	Λ.	3722	50	720	CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
		Į	Į	j	ļ	TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
		1				LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	A	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
		 		amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		GPFSKKDQYDVKAPAMFNIRNTGK/TLVART QGTQIASDGLKGLLFEVSLADLQNDEVAFRK FKLITEDVODKNCLTNFYGMDLTCDKICSMV
						EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK HNNQILKTSYA*HQQS/RQIQKKMMEIMT*EV QTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH DVFIRKVKMLENPGFER\MELRGGGSSS
1083	2433	A	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP WGAVQGSRAMSDLLLLLLDLTLLLLMLLGF AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
1084	2434	A	8950	156	318	GWLLT/ESCSISPKLCSIAVH*DNPAWF HYTPINTDTIENSENNKCW*GY*E\VGLIHHW WGGKRVQPFWKRVWQKRTLNLRV
1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK *TIYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR MFILAPFTATIKGKQLTCPLVEERIDY\MWYS HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW
1086	2436	Α	8962	868	1026	YSHKYY H*KILQVGRAQRAHXSRL*SQLLRRLRHESHL NPGARGCSEARLHRCTPAWTT
1087	2437	A	8985	58	330	LHVKHLGHFQLVFSEVICHCILMPVS*ELQRL *ERSVCAFHVCIQTYVCLQVYACMCVYYICM FVYSVYGCGLCTCVCMDVYICVCVQEFL
1088	2438	A	8989	394	404	N*KWILHVNVRIQSIFF/IKRNQK/INSHELKLD KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV YKDLSKLNRKTE/S*/VKKWVKDLSRYFIKE VISMENKHKKIFSTS
1089	2439	A	8991	60	329	MALTPESPSSFPGLAATGSSVPEPPGGPNATL NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*IHL
1090	2440	À	8996	2	351	SNITITLT+MKKYDNTFCW+GCGQIG/T/LIYC WQESKFIQAFWSKIQQYLA+ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI
1091	2441	A	8997	97	456	YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF
1092	2442	A	8999	548	811	SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG
1093	2443	A	9002	3	2745	ALGLQQPAGSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY TGPESRQEVLIRTDPESRQEIMCTGHESKQEV

SEQ ID NO: of NO: of NO: of nucleotide coided sequence No: of nucleotide coided sequence No: of nucleotide coided sequence No: of sequence Seque	IDFEFK VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE QLLEFV EMQGY KSSVI DGDKA DKDLK
eotide sequence ue	IDFEFK VTMVS GEIECS LLLED DLVRPL DTLLA FVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
sequence Sequence	TDFEFK VTMVS GEIECS LLLED DLVRPL DTLLA FVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 915 peptide sequence 916 peptide sequence 917 peptide sequence 918 peptide sequence 919 peptide sequence 910 peptide s	TDFEFK VTMVS GEIECS LLLED DLVRPL DTLLA FVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
amino acid residue of peptide sequence peptide sequence peptide sequence "T-Threonine, V-Valine, W-Tryptophan, Y-Tyrosine, X-Unknown, *-Stop codon, Y-Tyrosine,	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \text{\cong} possible nucleotide nu	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
peptide sequence	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
Sequence	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
PICTDPISKQEDSMCTHAEINQKLPVA LEALMCTNPEIKQEDPTNVGPEVKQQ DTEILKVARTHHVQAESYLVYNIMSS NTLEDELDQALPSQAFTYRPIRQRVYS CQDVTSTCLAVKEWFVYPGNPLRHPI QMTIPGGTPSLKILWLNQEPEIQVRRL CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLHAFIAQALCLQGKSTSQLV DYINPRAVQLGSLLVRGLTTLVLVNS, WKTSDFMPWNYPDGKLFHQK YLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGL+ HHWPSPLGLTPRREVGKTGLQLPQDC TGSHFVEATYKNPEDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGGG P*ITSSSCCLSAA/SHLSIQSPNMAGAR LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITEG 1097 2447 A 9032 716 357 ARSTOFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYR	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
LEALMCTNPEIKQEDPTNVGPEVKQQ DTEILKVARTHHVQAESYLVYNIMSS NTLEDELDQALPSQAFIYRPIRQRVYS CQDVTSTCLAVKEWFVYPGNPLRHPI QMTIPGGTPSLKILWLNQEPEIQVRRL CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLHAFIAQALCLQGKSTSQL* DYNNPRAVQLGSLLVRGLITT.VLVNS, WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGL+ HHWPSPLKHFVLAKKAITAIFDOT TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQTIGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRYY	VTMVS GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
DTEILKVARTHHVQAESYLVYNIMSS NTLEDELDQALPSQAFIYRPIRQRYYS CQDVTSTCLAVKEWFVYPGNPLRHPI QMTIPGGTPSLKILWLNQEPEIQVRRL CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLIIAFIAQALCLQGKSTSQL' DYINPRAVQLGSLLVRGLTTLVLVNS, WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLE HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDD TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWGKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVDFVPFGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQTIGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVFKPGCKEESTGTLFK GQHSETPSLKKKULAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRYY	GEIECS LLLED DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
NTLEDELDQALPSQAFIYRPIRQRVYS CQDVTSTCLAVKEWFVYPGNPLRHPI QMTIPGGTPSLKILWLNQEPEIQVRRL CFRI.SSSREELQAVESPFQALCCLLIYI DTLCLEDLHAFIAQALCLQGKSTSQLV DYINPRAVQLGSLLVRGLTTLVLVNS, WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVINSRVEDFVPPEGAGRTLFFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAAVSHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQITJQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKVLAGYSGMCL*SQVI EDCLSPPGGONCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	LLLED DLVRPL DTLLA FVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
CQDVTSTCLAVKEWFVYPGNPLRHPI QMTIPGGTPSLKILWLNQEPEIQVRRL CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLHAFIAQALCLQGKSTSQL' DYINPRAVQLGSLLVRGLTTLVLVNS, WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF- HHWPSPLGLTPRREVGKTGLQLPQDC VEVL/CRTK*ISAHQIPQPEGSRLQGLF- HHWPSPLGLTPRREVGKTGLQLPQDC Tegshfveatyknpeldrlateddlyi KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD VTTELDSWIDKFCTKSSTREITNSGSD LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGG P*ITSSSCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKULAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	DLVRPL DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE DLLEFV EMQGY KSSVI DGDKA DKDLK
QMTIPGGTPSLKILWLNQEPEIQVRRL CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLIHAFIAQALCLQGKSTSQLI DYINPRAVQLGSLLVRGLTITLVLVNS. WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLE- HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRFFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQTIGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGRNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	DTLLA LFVQV /NLQP ACGFP KGYA EGEQT LWV IVGAE JULEFV EMQGY KSSVI DGDKA DKDLK
CFNLSSSREELQAVESPFQALCCLLIYI DTLCLEDLHAFIAQALCLQGKSTSQLY DYINPRAVQLGSLLVRGLTTLVLVNS. WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF HHWPSPLGLTPRREVGKTGLQLPQDC HHWPSPLGLTPRREVGKTGLQLPQDC Tegshfyeatyknpeldriateddly KDKLSilgevlsrrhmkvaffgrtsc Namlwdkvlpsgighitncflsvegt YLMTegsdekksvktvnqlahalhm AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLFFALRPL LLHRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGAR! LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQTGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPRPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	FVQV NLQP ACGFP KGYA EGEQT LWV IVGAE QLLEFV EMQGY KSSVI OGDKA DKDLK
DTLCLEDLHAFIAQALCLQGKSTSQLV DYINPRAVQLGSLLVRGLTTLVLVNS. WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRFRLWPSCCCRV T*H\MAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSG NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	NLQP ACGFP KGYA EGEQT LWV IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
DYINPRAVQLGSLLVRGLTTLVLVNS, WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQTIGAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	ACGFP KGYA EGEQT LWV IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
WKTSDFMPWNVFDGKLFHQKYLQSE VEVL/CRTK*ISAHQIPQPEGSRLQGLF HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFD TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPEGAGRTLPFALRPI LLHRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	KGYA EGEQT LWV IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
VEVL/CRTK*ISAHQIPQPEGSRLQGLH HHWPSPLGLTPRREVGKTGLQLPQDC 1094 2444 A 9021 97 834 AREACRAKTDFPGRRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSS NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	EGEQT LWV IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
HHWPSPLGLTPRREVGKTGLQLPQDCC 1094 2444 A 9021 97 834 AREACRAKTDFPGRRFRLWPSCCCRV	LWV IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
1094 2444 A 9021 97 834 AREACRAKTDFPGRRFRLWPSCCCRV T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVPWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	IVGAE OLLEFV EMQGY KSSVI OGDKA DKDLK
T*HMAEPVSPLKHFVLAKKAITAIFDG TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD' 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLFFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	EMQGY KSSVI DGDKA DKDLK
TEGSHFVEATYKNPELDRIATEDDLVI KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD' 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLFFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	MQGY KSSVI OGDKA DKDLK
KDKLSIIGEVLSRRHMKVAFFGRTSSC NAMLWDKVLPSGIGHITNCFLSVEGT YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARD LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSFKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	KSŠVI OGDKA DKDLK
NAMLWDKVLPSGIGHITNCFLSVEGT. YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD' 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	OGDKA DKDLK
YLMTEGSDEKKSVKTVNQLAHALHM AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD' VTTELDSWIDKFCTKSSTREITNSGSD' LVLNSRVEDFVPPEGAGRTLPFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD PFFPNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKKLAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	DKDLK
AGCLVRVFWPKAKCALLRDDLVLVD VTTELDSWIDKFCTKSSTREITNSGSD 1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPL LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	
1095 2445 A 9022 1 537 LVLNSRVEDFVPPEGAGRTLPFALRPI LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	GPGTD
LLHRRARRSSALCPRPRSWGVSGGEG P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	
P*ITSSSCCLSAA/SHLSIQSPNMAGARI LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	AACW
LAKEKIEGCHICTSVTPGEPQVFLGKD DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	AGARE
DYVFDIDSQQEQIYIQCIEKLIEGCFEG FAYGQT\GAGKTYTMGTGFD 1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	URIRPQ
FAYGQT\GAGKTYTMGTGFD	
1096 2446 A 9029 1 285 FFFFNVCKSPKVPKPGCKEESTGTLFK GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	YNATV
GQHSETPSLKKK\LAGYSGMCL*SQVI EDCLSPGGGNCRES*SCPYTPAWITER 1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYRI	
1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYR	
1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVI LAHCNFRHAGFPPLSCLSLPNRWEYR	•
LAHCNFRHAGFPPLSCLSLPNRWEYR	
	ACLG
LPKCWDYRREPAASIIFQTTFFINSK	OWD TO
1098 2448 A 9038 230 652 KVVVMSCEDINISGSFYRNKLKYLAFI	
TNPSQGPYHLWVPSHIFWQTTCGRLPI	
G*AALDHLKVFDRIPLPYDKKKQMA\	
VVRPKP*RKFAYLGHWAQKVDWKYC TMGEKRKVYYQKICYQKK	AMIA
1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW	*C*VII
HSL/WK/TV+QFLKRLYLHLPHNSVIAF	
RKIKTCPQNSCTSMLINAIHNDQKWKJ	
1100 2450 A 9045 763 584 ROSLALSPRLECSGTISAHCRLCPLVFT	
SLTSSWDYRRPPPHPANFLYFK*RGF	LUCL
1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFA	7.1.11
FLL/FEMESLPVA/RVECSGTISAHCNLO	
DSPASAS*VAGITDMCRYTQLILFHAS	300
1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTI	KEIM
DEEDVLOECKAONRKLIEFLLKAECLE	
N*EEPPQDMDEKIRYKYPNISCELLTSI	
NDRLGEDESLLMKLYSFLLNDSPLNPI	
SKVLSILISRKPEQIVDFLKKKHDFVDL	
TSAIMDLLLRLLTCIEPPOPRODVLNW	ALAN III
RNL*HST*NVMDISKYVNLHWGLNKS	
LLLQCVLQWLNEEKIIQRLVEIVHPSQI	FKVQ
SLV	FKVQ HSLL*
1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSF	FKVQ HSLL*
WLYLQMYEIIKSPRFPIIKMTDITKCW*	FKVQ HSLL* EEDVS
AGMQI/H/CW\WCVNVGKFWEMS*YY	FKVQ HSLL* EEDVS
ST/PYDPAIPLLGIYL*ETRVYIHPKTCM	TKVQ HSLL* EEDVS WGLRE GC\GA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dence		914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine,
delice	Ì		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ	İ	peptide		/=possible nucleotide deletion, \=possible
		1		sequence .	İ	nucleotide insertion
						APFVLAVNC
1104	2454	Α	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI
					•	KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V
İ		}				KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
	İ	i				YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS
İ		İ	İ			RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR
ĺ					į	AWPCCPGWSAAWLTIVILAHYRRPGLERSCC
j						LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL
ļ						VLNS*TQGI
1106	2456	Α	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT
L		Ļ				HFPCDPAIPLLGICPED
1107	2457	Α	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT*
						AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA
		ļ	1			VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL
	i	}				I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA
1		1]			PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP
<u></u>	0.150		0000	7.10		ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	Α	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP
1						GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA
Į.		ļ	ļ		ļ	SAFPPAERSRGHRRASL*RARWSAAVPRRSA
1						GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG
						ORPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL
1105	2437	Α	7077	1233	1423	GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	A	9103	242	70	EEOFFFFAVGMFP*VDFLAPASGELWDRLRLT
	2.00	**	7.05	2.5	'	CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP
						SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS
						LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	Α	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP
ł						AAAGDPASLDFAQCLGYYGYSKFGNNNNYM
						NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT
						PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT
					ĺ	PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM
						DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG
						VMPPAQLTTINQSQLSAQLGLNLGGASMPHT
						SPSPPASKSATPSPSSSINEEDADEANRAIGEK
L	2462		0.00	2452	0051	RAAPDSGKKPKTPKK
1113	2463	Α	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F
	1					SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH
						VGQAGHEPLTSGDPPASASQSAGITGVSHQA
						WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS
1114	2464	•	0122	163	277	LLKCWDY
1114	2464	Α	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS
						SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465		9124	553	981	
1112	2403	A	7124	ددر	701	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM
						TOTOSLSFLLGSSASLDCGFSMAPGLDLISVE
		1			[WRLQHKGRGRGDLHLPDHHLSVPSSADHPA
1						QQPSQFNGRNLYFLPLFR
1116	2466	Ā	9135	48 .	410	SASHEPAEHDGGADSLSASOPPRPAGRPAGA
''''	2400	^	21.23	TO .	710	QHVHVPPWTDVLAGQDRRAPTAGDGAPWP
						APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ
						PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA
]						GAGGP*GSPAGRACGAAGCRPRPPRPAASSA
						*NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS
					<u> </u>	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTITVELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP
					ŀ	LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*
1118	2468	Α	9154	471	2	GILTKLSAFLTIPRLQPHLIAALSPSS AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	2	3187	ACPRLARRRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDA VLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSVLRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQGULRG\ WKARRAITTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEDEDDDGGEEAPAPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1120 1121	2470 2471	A	9163 9166	124 272	207 523	PPRACRPCPRACPCPPT*KCSQPVSWPC PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
				-		V/CSHITDSLKFIGKGWVGMVTHACNPGTLG
1122	2472	С	9170	442	236	G*GGWIA*VREFETSLGNM MNRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VIFTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1124	2474	A	9173	peptide sequence	374	nucleotide insertion GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
1125	2475	A	9179	704	188	TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA IAHRGGRHDPPENTLGAIR/QGS**WSNRR ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
1123	2473	A	9179	704		CSKPPKETGELENAESGGDGGRRGGKQDNV AWWRRMQKG\DFPWDDEDFPQSGPFGQA LPMGFFYLYFRDPGREITWKHFVQYYLARGL VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	Α	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM EYMAESTDRSPGHILCCECGVPISPNPAQY\CV ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC L\LDGVPVALKKVQIFDLMDAKARADCIKEID LLKQLNHPNVIKYYASFIEDNELNIVLELADA GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS ALEHMHSRRVMHRDIKPANVFITATGVVKLG DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD NG
1129	2479	Α	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT RATT\KIRVVATITRARIEDMRHSATALTRPD ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE PLAFLRALELLFAIFAFATCGGYSGGLRLSVD CVNKTESNLSIDIAFAYPFRLHQVTFEGIPTCE GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL AATGVYIFFONKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK TQPVEATDDAFWDQFWADTATSVQDVFALV PAAEIRAVREESPSNLATLCYKAVEKLVQGA ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD WRGFFWSTVPGAGRGGQGEEDDEHARPLAE SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME LLKLLTCFSEAMYLPPAPESWQH/RTHWFSS FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY NHLY
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS NVYFIV
1134	2484	A	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK RADGAAPAGEGEGVTLQGNITLLKGVAVIVV AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC GVFSIVGALCYAELGTTISKSGGDYAYMLDV YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC YSVKAATRVQDAFAAAKLLALALIILLGFVQI GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII GMIWLRHRKPELERPIKVNLALPVFFILACLF LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ RDRLPPAYFCRPVVCVVTALDVG\SPESQEM
						DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKNPRNNL
1136	2486	A	9223	3	983	RSLLGERVDENTEENHCGETSSQIPDDTLNK RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG DLVFAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD SEAPEANPADGSDADEDDEG\RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK MSVSKRARKASSDLDQASVSPSEEENSESSSE SSKTSDQDFTPEKKAAVRAPRRGPLGGRKKK APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK PKPERPPSSSSSD
1137	2487	A	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA
1138	2488	A	9231	207	443	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TIPQGKPETIFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
						VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL
1140	2490	A	9238	248	328	MAQGNNYGQTSNGVADESPNMLVYRKV
1141	2491	A	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN
1142	2492	A	9245	157	466	HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ
1143	2493	A	9247	264	115	GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG ARDSTSIIRMGPEIPPP

SEQ ID SEQ ID Met SEQ Predicted Predicted end Amino 8	cid sequence (A=Alanine C=Cysteine,
	rtic Acid, E=Glutamic Acid,
	vlalanine, G=Glycine, H=Histidine,
eotide seq- USSN location corresponding I=Isoleu	cine, K=Lysine, L=Leucine,
seq- uence 09/496 correspondi to last amino M=Meth	nionine, N=Asparagine, P=Proline,
	mine, R=Arginine, S=Serine,
	onine, V=Valine, W=Tryptophan,
	sine, X=Unknown, *=Stop codon,
	le nucleotide deletion, ⊢possible
	de insertion
	RLSEMSFSLNFTLPANTTSSPVT\DCGP
	AGIPLLVATALLVALLFTLIHRRRSSIE
, , , , , , , , , , , , , , , , , , , ,	DRPCEISEIDDNPKISENPRRSPTHEKN
, , , , , , , , , , , , , , , , , , ,	QEAHIYVKTVAGSEEPVHDRYRPTIEM
ERRR	
1145 2495 A 9264 175 411 METIW	TYQFRLIEIGDSTVGKSCLLHRFTQGRF
	PACDPTVGVDFFSRLLEIEPGKRIKLLL
	GQERFISIT
	EGREGIKSQPKMEPHSVT\RLECSGMI
1 1 1 1 1 1 1	LNLPGTSDSPASASR/VAGTTGMRHHA
	FLVETGF
1 1 1 1 1 1	RGEEEKEEEEEEGWVNGMENSHPP
, , , , , , , , , , , , , , , , , , , ,	QPPPQPGPSGERRNHHWRSYKLMIDP
	HHKLYRYDGQHFSLAMSSNRPVEIVE
	GIWTKNKE\LELSVPKFKIDEFYVDQV
	TFAKLNDNIRENFLRDMCKKYGEVEE
	JPKTKKHLGIAKVVFATVRGAKDAVQ
	SVMGNIIHVELDTKGETRMRFYEL\LV
	PQTLPVGELDAVSPIVNETLQLSDALK
1 1 1 1 1 1	GLSAGCGSGSSSVTPNSGGTPFSQDTA
	LDTPNSYG/QGTPLTPRLGTPFSQDSSY
	PSYLFSQDPAVTFKARRHESKFTDAY
	HHYVHNSPAVTAVAGATAAFRGSSD
	/GGTGGSSGPPFKAQPQDSATFAHTPP
	PAPGFR
	ADTMPGVGLLVSHFSTLVSRQRCPNY
	LTDVSIFLLLEVSGDPELQPVLAGLFL
	TVLGNLLILAISPDSHLHTPMYFFFSN
	\AGFTSTTVPK\MIVDI\QSRSRVISYAG
	SLFAIFGGTEE\NMLLSVMAYDRFVAI
	HSAIMNPCFCAFLVLLSFFFLSLLDSQL
	LQFTIIKNVEISNFVCDPSQLLKFACSD
	IYFHKDPERQLVLAGLFLSMCLVTVL
	DVSPDSHLPTPMYFFLSNLSLPDIGFT KMIVDIQSHGRVIFYAGCLTQMSLFAIF
	RHAPECDGL
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	KDIFIGWGTIHLFRKPQRSFFGKLLRE
	ADRSMGRYMLFGVINLICTGFLLMWC ALIASYTYLTIFDLFSLMTCLISYWYTL
, , , , , , , , , , , , , , , , , , , ,	/YSFGFERLEVLAVFASTVLAQLGALF
	ERFLEQPEIHTGRLLVGTFVALCFNLF
TAG CIT	NKPFAYVSEAASTSWLQEHVADLSR
	PGLSSIFLPRMNPFVLIDLAGAFALCIT
	/TRAGVQWCSLGSLQPRTPGLLRSSCL
1130 2300 A 9308 797 693 DK3130	TIVO A Á M COTOOPÁLK I LOTTVOOCT
	PVLWKWSKPTR\TAKEPPQTQQRAG
1 (1 1 1) [1 1 1	` ' '
FLICTI	PPCQWSRMASEGPNIPCPGARHSDKQ
	TODA VVII DOCAVII NII VA PERGOPA OVEN
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	TPPAVVLPPSAVLNLVNTFSSFPQVEV GPRKGRLAVTIPFFGLS/LPKYMDHRR
1 1 1 1 1 1 1	1
, , , , , , , , , , , , , , , , , , ,	EIFFVFLAETGFHRASQAGPDLPTS/S/I
	FPKCWEYRSEPQCLPGCLSFSGILLDL
GTNVS	
	RPGFRSPARSSRPCPVLTSLLPPFPSPSP
i i i i i i PADDI.	VKAGRDRKDPQVR/ERRLRPNPGRLG
. GPR\PR	PARARS/CHQPRLTRVCPRSPPPEARA
. GPR\PR	PARARS/CHQPRLTRVCPRSPPPEARA PARGRGAPKRNRPRTDTRAPRGSSAR

CEO ID	CEO ID	Mat	Lego	Predicted	D. 2: 4-2 - 2	I Amin and an arrangement of the control of the con
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l liou	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	4000	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	,	/=possible nucleotide deletion, \=possible
			1	sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
		1	ĺ		ĺ	PT
1155	2505	A	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
		j	J		j	EAAQRHCRASVSILRMRRPGQGSSRPARVPL
		ŀ				RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGGSHPAGAQPGGRAIGAGWQS
		1	:			KEPLWEGLQRSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
						LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
		1			!	VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
		l				SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
		ĺ			ĺ	RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
		1				LMEAGLPQKQAERADELFEAGLVIYVKLDER
]	Ì		j	VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
						KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
					ļ	DHTDQELREEIHKANVERVVHDVSQEATIEKI
		ł	ļ		ł	RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
		ļ				EAELPIMSQLTEIETCVEC
1161	2511	Α	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
		ľ				AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
		1				DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
						QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
		.				GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
		i				TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
		1				VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
		ĺ				VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
ļ		İ				GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
		l				FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
		ļ				WIDGPKDTSVEDALERTYVPCLKTLQEEVME
						AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
						G
1163	2513	Α	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
		!	!			LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
			1			AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
		<u> </u>				FSSYQRNNPDLILNDTIMPNIK
1164	2514	Α	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
		•	[HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
			1			FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
		1				VAGASVGAGVWARNPRYRTEGEACVEFKA
ļ		}	1			MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
						VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
						SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
						SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
ł		ł				VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
						YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
						AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
1166	2515		0262	647	001	LPLQNKDRGSWPASRGSPRLL
1165	2515	A.	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
						VPEGVRLADGPGHCKGRVEVKHQNQWYTV
						CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
j			}			TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
1166	2516	-	0262	201	202	LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	Α	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
1167	2517		9368	707	1087	GAISAHLAHCNLCLPGSSDSPASAFQVAS AVLTPCLSPCSPSRIPRPSRPYPGRRSLSHTPP
			· UANK	. /11/	: IIIX/	. AULIPET SPESPSPIPPPNSPVVDEPPST SUTPP

SEQ ID SEQ I			10 11 1	D 11 1	14
NO: of NO: o	•	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl- peptid		in ID NO.	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	·	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	peptide		/=possible nucleotide deletion, \=possible
		ļ	sequence		nucleotide insertion
		1			PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
		.	}	[ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
			!		PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168 2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS
, ,	Ì	ļ	ļ	J	PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID
		}	Į.		LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS
1 1	1	1			QQSILAGLVVVATTGMIGSPLECLFGELGGRA
			1		DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL
		1			FNAMQASSEH
1169 2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN
1					NSARRMEAMASGSNWLSGVNVVLVMAYWS
	İ				LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH
					NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170 2520	Α	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR
ł	1	İ		Ì	ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF
					TNETWQARTGEPLPDHLVLLMWSLIVSLYPL
				1	GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS
		1			AAILFGFSRKAGSFEMIMLGRLASWGVNAGV
					SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA
	İ			•	LGIVMGQVVGLSTTAATGLRGL\AGELEELEE ERAACQGCRARRPWELFQHRALRRQVTSLV
}	1	1			VLGSAMELCGNDSVYAYASSVFRKAGVPEA
					KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL
					WGGTPRSFALNQFTLQKKKK
1171 2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA
2321	' '	7501	-	712	LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE
	1				TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV
					EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG
[SKRLFFWMQEP
1172 2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKET\KAAMGTQ
					CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL
					LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL
					ASFNEVGNTALIVLESY
1173 2523	Α	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ
		1			QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI*
]			KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV
L					IRPPISFSKINNGP
1174 2524	Α	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ
		J			RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF
			ļ		ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF
		\			LGML
1175 2525	Α	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT
					AKMNKWSKGKVRDKLNNLVLFDTATYDKL
) ·		CKEVPNYKLITLAVVSERLKIPGSLARAALHE
1176	\rightarrow	0400		200	LLSRGLI*LVIQHIAQVIY
1176 2526	Α	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY
	1]			GERGYAQNGDF*DAQLDDYSFSCYSHAQVN
	1				GAPNSLTRAYDDP*VKISGLECQKVGALVEV
1177 000		0000	<u> </u>	400	KCLNL
1177 2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV
1	1				YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD
					FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF
1170		10410	140	406	DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178 2528	Α	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP
l I	- 1	1			ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL
ļ	J				
1170	 	0420	1450	1/55	ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179 2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG
1179 2529	A	9420	1450	1655	

SEQ ID Not of modified N		T					
moclocide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Coation Sequence March Sequence March Sequence March Sequence March Sequence March Sequence Seq	•		noa	1			
Sequence	1	1	ł		1		
United Part							
amino acid of peptide residue of peptide residue of peptide residue of peptide sequence y=1yrosine, x=1option, y=1yrosine, x=1option, y=1op		uence					
residue of peptide sequence	uence		İ	914	1 –		
	1						
1180						sequence	
1180	1	ł		}		ł	
SRFSRSTRPVPPPRADPPARQKLTGVLHAPILIK L	1100		<u> </u>	1		0.75	
1181	1180	2530	A	9422	170	3/3	
1181		i					
	1101	2521	<u> </u>	0426		-074	
	1101	2551	A	9430	4	2/4	
1182	ł	1		ŀ	i	ĺ	
SMILLE*MGAGDEKISAMGKARVDHRELYLÖL LYPTEDYKLIFRARH	1100	2522	<u> </u>	0442	,	240	
LYPIEDYKLITRARH	1182	2532	A	9442	3	240	
1183	Ì			1		İ	
KGCSGWAPWISLQCQHFGRPRWADHLRSGV RDOPPGGYSKTTEPKYQKLAGSAHL*S9*L1 ERMRWKNRLNPGGRSCSEPRWHCTPGWAT ERG	1102	2622	<u> </u>	0444	204		
RDOPGQYSKTTFLPKIQKLAGHSGAHL-S*LL	1103	2533	A	9444	384	3	
ERMKWİNRLNPGGRSCSEPRWHHCTPGWAT	!	ļ					
BEG		Ì		}			
1184		1		ĺ			1
RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV	1104	2524		04/0	701	(55	
HITCNPSTL.GGR.AGWIV*AQEFET	1184	2534	A	9462	391	633	
1185	1	į				1	
WWWGWEGWERVALKLSSGPAGPLACWVAK KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG WSAVVQFWLTAASNSCFISLLSSWDYRCA WSAVVQFWLTAASNSCFISLLSSWDYRCA HIPQLHTKTHYVPTRMYNK!*QIDNSKPWQR GG*TGILTHCW*ESKLVQPLWK!VWHYQ SRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN SRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN SRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN SRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP NPASPHPEAPQEPWDSASGSVGSFSLGRGAK ASS*VPGKGRGPRQGSELLAETILELFLALAN SRGAMAPEALKILSAITQPMVEEALAGLYRAC *PYLTNILAGMKKGLCLGSTEQAHTIGI TIMB SASS	1105	2525		0467	215	566	
KKSLSLSGPVYPSEKGAGLYVP*DRVSLCHPG	1165	2535	A	9407	213	300	
WSAVVQFWLTAASNSCFSLLSSWDYRCA	ļ					!	
1186					1	į	
GG*TGILTHCW*ESKLVQPLWKIVWHYQ	1106	2526		0469	775	452	
1187	1100	2330	^	9400	2/3	432	CC*TCII THOW*ECKI VODI WKIVWHYO
SSRGAEPCLSNCIHSPAPRKQRIMGDSDO*STP NPASPHPEAPQEPWDSASGSVGSFSLGRQAK ASS*VPGKGRGPRQGSELLAFTILELLFLALAN S	1107	2527	<u> </u>	0460	200		
NPASPHPEAPOEPWDSASGSVGSFSLGRGAK	1107	2337	A	9409	300	•	
ASS*VPGKGRGPRQGSELLAETILELFLALAN S S	ŀ		1				
S	l			1		i !	
1188]]		J			1
GRLMANPEALKILSAITQPMVEEAIAGLYRAC	1188	2538	Δ	9471	124	397	1 T
1189	1	2330	'	''''	127	""	
1189	ļ			1			
PSLLKIQKISWAWWRAPVVPATWEAEAEEW R	1189	2539	A	9480	584	769	
R	,	2557	,	1	501	'**	
1190]]		ļ			
PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK	1190	2540	A	9483	463	86	L
GASSCRRRCNPVLAARKAGSPRSHSTRENC RRSRCPDTAHRRRRGRRRNPSCVRSPRWR 1191 2541 A 9489 I LADALCLSAAATGAVRPGARAQPSTRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEENTS 1193 2543 A 9509 186 I IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP					100		
RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR	1						
1191 2541 A 9489 I 411 LADALCLSAAATGAVRPGARAQPSTRRRLSP SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEELTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	l						
SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEELTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRPPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1191	2541	Α	9489	1	411	
ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP		l] .				SVRVCCRAAAASNLLYSSCLORHSERASEEG
MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK 1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTTVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP		Į.	١ .				ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
KEELTAVNVK			ŀ				
1192 2542 A 9497 389 161 VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI* CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1	}	ŀ				
CEEDERKMAREFLAEFMSTYVMMNIHMIVE KDTYSDHEEINTS 1193 2543 A 9509 186 1 IAKSQ*KRWQRSGAMETLKHGWWECKLVQF FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM 1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1192	2542	Α	9497	389	161	
KDTYSDHEEINTS	ĺ	1	1			1	
1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFFR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRPPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP		}	1			}	
1194 2544 A 9512 58 433 PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRPPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1193	2543	Α	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP		}	ļ	1		1	FGKTFVNVN+S+TYVYPCDKIILLLGLYPTEM
LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1194	2544	A	9512	58	433	
SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ 1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP			l				
1195 2545 A 9515 595 1223 GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1	1	l	1			
AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	<u> </u>	1	l	1			
AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1195	2545	Α	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL* PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP		1		1			AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC
PLHLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP	1		1			[PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
APGSGPCGATARPSRGGRAGGSRARRPIPPGP			l	j	1		
		-	l				
GTRRTPSGCQNPAASGG		1	I]		[
	L			<u> </u>		L	GTRRTPSGCQNPAASGG

NO. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Seq	NO: of		hod	ID NO:		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence		, - •					
Uniform Section Sect			l	1			
amino acid residue of peptide residue of peptide residue of peptide sequence per per per per per per per per per pe		uence					
Peptide	uence	-	l	914			
peptide	-		1	İ			
1196	Į	,				sequence	
1196							, ,
ARRVAPGWDLHTFYLPRINSRTPH.**EPHA	1196	2546	A	9518		468	
1197 2547 A 9521 289 448	1		' '	70.0		'''	
HAMRHLTGMTSMAIREL*ADSRFQVQRARYE	1		ł	1	į		
APNWKYKYGY-IPVDMLC	1197	2547	Α	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
1198	ļ	1				Ī	
VKQEPNICHL*ETHFFFRDTYRLKEREQKKRK SSYS							
1199 2549 A 9546 1785 1943 GGREKESIL:TNAGWQRNSFIGPPKSIPWAA	1198	2548	Α	9524	204	1	
199				İ		1	1 3
1200 2550 A 9548 186				<u> </u>			
1200	1199	2549	Α	9546	1785	1943	
			ļ			<u></u>	
1201 2551 A 9549 591 2 SSVVEPRGPRSSLPPLDSTFFCGSSPNWTGGC GSCPSGET-VSPGSEQRKYSNSNVIMHETSQ YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS KEKIWTQEMLLQVNDQSLRLIDISGELEDF PLFTVQRSQTVLNQLRYSVSNSVIMHETSQ YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS KEKIWTQEMLLQVNDQSLRLIDISGEELDF PLFTVQRSQTVLNQLRYSVSVLLIVCQDSEQSK PDVHFPHCDEVEAELVHEYMESALTDCRLGK AMRP	1200	2550	A	9548	186]]	
	1001	0551		0540		<u> </u>	
1202	1201	2551	A	9549	391	2	
REKIWTQEMLLQVNDQSIRLLDIESQEELEDF PLYTVQRSQTVLNQLRYPSVLLLVCQDSEQSK PDVHFHCDEVEALLVHEYMESALTDCRLGK AMRP	1		i			}	
PI-PITVQRSQTVLNQLRYPSVLLLVCQDSEQSK]	İ			
PDVHFFHCDEVEAËLVHEYMESALTDCRLGK			1	}		•	
1202 2552 A 9552 428	1	1	ļ	1			
LDCERPPQGPLPSLPELAKTSYSDLTGLATED	i	ł	l			i	
LDCERPPQGPLPSLPELAKTSYSDLTGLATED	1202	2552	Α	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
1203 2553 A 9568 517 738 RRKPERKOKO*RVREGKQYRQRDKMKEWG EKEKRREKGERERKMRHRERKGESGQRD TMENWRVERLTEKER 1204 2554 A 9573 83 415 EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD DGWDLSDAHVVCQKLGGGVAFNATVSAHFG EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ HDCRHKEDAGVICSEFTALR 1205 2555 A 9577 64 424 ARGSCPTRPRTANGRMGETKDAPQMLVTFK DVAVTFFREWRQLVLVHRTLYR*GMLETC GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE VSHSPCAGDMRELFTREATLTPHPYNNGA 1206 2556 A 9584 38 476 TLGAVLFSEVESSTSHSGQLGRQNRIPKL SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV NPQPLSTPSWQIETKYSTKVLTGNWMEERRK GLPYKHLITHOEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL 1207 2557 A 9586 2 412 LRSSPAALLRALCITTVTGTALALRSRVATTN PDGGCRNVLRPEY YNGLOEAGSWGALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGFTRFFL FGILESICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 SISLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIVQ RIRDHDLLDKRKTVTALKAGEGRALLGLAM MVCSIMM*FLIGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRLEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRLEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRLEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS****WSPKNSHVGKDS 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRLEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS 1210 2560 A 9618 384 2 SLHDMLMLAEQQGCGPGDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS 1210 2560 A 9618 384 2 SLHDMLMLAEQQGCGPGDDI KVQVKNNDLGQATNNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS 1210 1210 1210 1210 1210 1210 1210 1210 1210 1210 1210 1210 1210	1			1		1	
SKPRATPLFCSLHTF							*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
1203	ĺ		l	İ			
EKEKRREKGEREBREMRHRERKGESGQRD							
TMENWRVERLTEKER	1203	2553	A	9568	517	738	
1204	İ	İ		ĺ		,	•
DGWDLSDAHVVCQKLGCGVAFNATVSAHFG EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ HDCRHKEDAGVICSEFTALR 1205	1204	2554	<u> </u>	0573	92	415	
EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ HDCRHKEDAGVICSEFTALR 1205 2555	1204	2334	^	3313	65	413	
HDCRHKEDAGVICSEFTALR	ľ		1				
1205					ļ		
DVAVTFFREEWRQLVLVHRTLYR*GMLETC GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE VSHSPCAGDMRELFTREATLTPHPYNNGA 1206 2556 A 9584 38 476 TLGAVLFSEVSKESSTSHSGQLGRQNRHPKL SNFITPSSPRLK*TASSQRNLGQILNMFLTAV NPQPLSTPSWQIETKYSTKVLTGNWMEERK GLPYKHLITHHQEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL 1207 2557 A 9586 2 412 LRSSPAALLRALCITTVTGTALALRSRVATTN PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGGRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1205	2555	. A	9577	64	424	
VSHSPCAGDMRELFTREATLTPHPYNNGA		!					DVAVTFFREEWRQLVLVHRTLYR*GMLETC
1206 2556 A 9584 38 476 TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV NPQPLSTPSWQIETKYSTKVLTGNWMEERRK GLPYKHLITHHQEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL LRSSPAALLRALCITTVTGTALALRSRVATTN PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	}						
SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV NPQPLSTPSWQIETKYSTKVLTGNWMEERRK GLPYKHLITHHQEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL 1207 2557 A 9586 2 412 LRSSPAALRALCITTVTGTALALRSRVATIN PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK		•					
NPQPLSTPSWQIETKYSTKVLTGNWMEERRK GLPYKHLITHHQEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL	1206	2556	A	9584	38	476	
GLPYKHLITHHQEPPHRYLISTYDDHYNRHG YNPGLPPLRTWNGQKLLWL 1207 2557 A 9586 2 412 LRSSPAALLRALCITTVTGTALALRSRVATTN PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK						ĺ	
YNPGLPPLRTWNGQKLLWL		i		l			
1207 2557 A 9586 2 412 LRSSPAALLRALCITTVTGTALALRSRVATTN PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK			l	1			
PDGCRNVLRPKYYRLCDKAESWGIALETVPT GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1207	2557	A	9586	2	412	
GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	120,	٠	''	7500	_	*12	
THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK		İ	.				
FGILFSICFS 1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGGRGRWIA*AQK FADAWADAW 1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK]			
1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	L	1	l				
1209 2559 A 9611 148 558 KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1208	2558	Α	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK			<u> </u>				
RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1209	2559	A	9611	148	558	
MVCSIMM*FLLGITLLRSYMQSVWTRESQCT LLNASITETFNC 1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	I		ĺ				
1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1		1				
1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAVNTQNTAWSNA DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	l		1	1			
DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1210	2560	_	0610	294		
KVQVKNNDLGLQATINNEANWIAHQDDFNW LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1210	الالالا	^ ^	8105	J04		
LLAELNTCQRQETADS***WSPKNSHVGKDS GELSAK	1		 				
GELSAK	1		l		1	İ	
			}				
	1211	2561	A	9620	316	610	

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
						GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST EENFGEKLHDIGFGNGFLDKT*KAQATKAKI DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI NHLPETERNLLEHGLMYIRLNAAFCSLVAHS LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV EEHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL GALGGRGGRALGGSRWPPPLPGETLFSGCKH RRRRGSDAAPGEEAGT
1218	2568	A	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL
1220	2570	Α	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K
1221	2571	Α .	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS
1222	2572	A	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE
1224	2574	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE GRRPRAVKVYTINLESQYLLIQGVPAVGVMK ELVERFALYGAIEQYNALDEYPAEDFTEVYLI KFMNLQSARTAKRKMDEQSFFGGLLHVCYA PEFETVEETRKKLQMRKAYVVKTTENKDHY VIKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	""	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			!	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	1	/=possible nucleotide deletion, \=possible
1	·			sequence	ļ	nucleotide insertion
1225	2575	A	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
				_		TDDVHINDISLQDYIAGVRLILL
1226	2576	A	9713	82	492	OGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT
						ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
			!			ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
						WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
1		1	i			MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
			!			LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
1						NRKKMKDCQLRKQQNENVSRAVCALLNSGG
			[]		ĺ	GVIKAEVENKGYSYKKDGIGLDLENSFSNML
						PFVPNFLDFMQNGNYF
1228	2578	Α	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
j						TLVLQKSDVEAVF
1229	2579	A	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
,						GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
			i			QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
						PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
						NETDLAGPMVFCLAFGATLLLAGKIQFGYVY
			į (GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
						GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
						WCSFSASKIFISALAMEGQQLLVAYPCALLYG
						VFALISVF
1230	2580	Α	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
			j			HFSPERPFMDYFDGVLMFVDISGKCKRDVCL
						MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
1	,					ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
						LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
1				•		LRCGWSPAEELNYTVPGPGPAGEASPRQCRR
						YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
						RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ
						SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
			. '			NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
1						WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL
!] .	•		LVLAGVAYALPHWRWLQFTVALPNFFFLLY
ŀ						YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG
1232	2582	A	9753	164	517	KSLPASL PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
1232	2302	Α	2133	104	217	YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
			[LLOTVFVSCLLLSAPVFGYLGDRHSRKATMS
						FGILLWSGAGLSSSFISPRYSWLF
1233	2583	A	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
, 233	2000		''''	2.5	717	IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
						KRAYKSYVRALPLLKKMGINSILLRKSIGALE
]			VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
		·				FHOLV
1234	2584	A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVOP
	2007			<i>''</i>	100	CSRCGYGVYPAEKISCIDOIWHKACFHCEVC
			ļ .			KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
1					İ	VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
						FHWD
1235	2585	Α	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA
						SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
						SEDPGPNWKLKVTSGLLIRGOTGEEMTRDGP
						ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
		·	[[!	CARPGHSARNNTDKNLPHTAIILVTSNTYTTI
					•	KINFQAGRSGSCL
1236	2586	Α	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC
	لستسا		لــــــــ		<u> </u>	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuc l -	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		·	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
						LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
						DGFVIVYDISDRSSFAFAKALI
1237	2587	Α	9793	266	515	NILAUYFPFPRLFLLRDSQSNPKAFALTLCHH
			ŀ			QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
	0.000	L	0000		0.55	RYKPTRPVCITQFQGCS
1238	2588	A	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
		ł				FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT TDEDSTRFOIINEASKVPLLAEIYGIEGNIFRLK
						INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
				ļ		GSLILADGKGDLKC
1239	2589	A	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
						MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
ı						GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
]	1	1	<u> </u>		PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
			<u> </u>			DIKFLTKGDNNEGDDRGSYK
1240	2590	Α	9819	3	305	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
						PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
						YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF DPNSEEHG
1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
1241	2391	^	9634	041	1209	LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
					ł	GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
		ł	1	l		SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
		<u> </u>				LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
		1				DGSVKCTHQPCLIKWISERGCWSCELCYYKY
		,	<u> </u>			HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
	}					LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
	ŀ	ŀ	Ì	l		MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP
1243	2593	A	9846	198	411	WRISHHAGKMPVMKGLLAPONTFLDTIATRF
1243	2373	1	7040	170	711	DGTHSNFILANAOVAKGFPIVYCSDGFCELAG
			1			FARTEVMO
1244	2594	Α	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
						PANGAGGPGGASARKLSTFLGVVVPTVLSMF
		1				SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA
		İ				LTVLSVCAIATNGAVQGGGAYCILQHRWTG
						VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
1045	2505	<u> </u>	0040	633	1620	NVCGCAVSLLGLVESVLDVFGA
1245	2595	A	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
	1					SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
	1	[[AMSLTEDVLAAALADHLPEDKWSAEKRRPL
						KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
			1			YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
		İ	İ			RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
	}	}	1	l		SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
	1					NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
				1		EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
	!	1	1	1		LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA
1246	2596	A	9850	114	464	PPOLGAORVREPRHPDVRAPLRVTSPGLRSRS
1240	2,50	^	0000	***	107	ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
						QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
						RRERPAGADSLSWGAGPRISSYV
1247	2597	Α	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
		l	1			HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
•] .		GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
			<u></u>			LYKVFKYSSRCTS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RVDDFVYSKGGKDAGGADVSLACRRQSIPEE
						FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	Α	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARIIAPRRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGLLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	A	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA
1251	2601	A	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	Α	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	Α	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL
1255	2605	Α	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT
1256	2606	Α	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	Α	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence.	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH QRRGPSCGASGDPQCVGSPHPQRARPLLARP GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS YLGECGSSSYVTGAACISPVLRCREWFEAGLP WPYERGFLLHQKIALSRYATALEDTVDTSRL FRSRSLREFEEALFCHTKSFPISWDAYWDRND PLRDVDEAAVPVLCICSADDPVCGPPDHITLTT ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR ERPAARSGPEMRVRYPVVAAVLAPYLALSQD PMYKSSASGQGASGSYNHVREEMLIKAGGA MSRRVVRQSKFRHVFGQAAKADQAYEDIRV SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL PLAK
1259	2609	Α	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL PPRPGRSHRKRKLVSTK
1260	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE KKQKASQNLVVLAREDAGAEKIFRSNGVQLL QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	ì	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF PTRVDHNGALLAFSPPPPQRQRRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY WTREGLAWQRADRPHCLYA
1262	2612	Α	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	Ā	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL
1264	2614	Α	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY
1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	Α	10005	2	209 ·	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide sequence	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1269	2619	A	10010	245	688	TKVSLGLA FGMLKNKGHSSKKDNLAVNAVALQDHILHD LQLRNLSVADHSKTQVQKKENKSLKRDTKAI IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS QPKLDRTSSFRQILPRFRSADHDRYRGWSMW DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP FLSGAEVSQSCRRRGGRA
1272	2622	Ā	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	3750	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT
12/7	2024	Α.	10017		J 130	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT ANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTYGLMGAVNVAK GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VAKGAIQGGI.DTTKSVLTGTKDAVSTGLTGA VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQTGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL QDCFRL
1275	2625	Α	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDQ
1277	2627	A	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	Ā	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	Α	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	IIOu	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
		ĺ		location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	j	USSN			
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł	ŀ	1	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l	t	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ	İ		peptide		/=possible nucleotide deletion, \=possible
	l	1		sequence		nucleotide insertion
					1	YGYVTADFISQSSSASPGGVDYILHGSTVTFQ
ĺ	ĺ	ĺ	1			HGONLSFINISIIDDNESEFEEPIEILLTGATGG
	l					AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA
	İ	ļ				NPNSTMILSLVLERTGGLLGEIQVNWETVGPN
Ì	ŀ	l				SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII
ł	ł	ł	ł		į	LTIYPHEEIEVEETFIIKLHLVKGEAKLDSRAK
1	1	}				DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL
ļ	ŀ	ł				
		<u> </u>				ALEGPLLITFFVRRVKGTFGEIM
1283	2633	Α	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT
						VPSDHLPNLYGFSALHAVHLHQWTLGYPAM
<u></u>	L	<u> </u>	<u></u>		<u> </u>	HLXRS
1284	2634	A	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP
1	l [']	l		[LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV
		l				TSLMPVLLFPLFQILDSRQVCVQYMKDTNML
1		ľ			ļ	FLGGLIVAVAVERWNLHKRIALRTLLWVGA
l	ì					KPARLMLGFMGVTALLSMVISNTATTAMMV
İ		l	1			PIVEAILQQMEATSAATEAGLELVDKGKAKE
		ļ				LP
1285	2635	A	10092	290	728	KOSTRPDVMTLYPLHWOEEMSGESVVSSAVP
1203	2033	^	10092	290	126	AAATRTTSFKGTSPSSKYVKLNVGGALYYTT
}	j	İ				
		ļ			ļ	MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL
	1					IDRCGKHFGTILNYLRDGAVPLPESRREIEELL
						AEAKYYLVQGLVEECQAALQV
1286	2636	Α	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS
ł	l	l	1		}	QRGSDAAGTMGCCTGRCSLICLCALQLVSAL
		ļ				ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ
1	l	ľ				YRPRYIMVYTVWTALWVTWNVFIICFYLEVG
ł	1					GLSKDTDLMTFNISVHRSWWREHGPGCVRR
ļ		1	1			VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI
]	ŀ				HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL
						GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS
1200	2030	l ''	1010,	1	476	AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL
ļ	l	j	1	ļ		RWRRCRSPRSEPRSQESGTDTATVLDMATD
1	1	l				SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM
	1	1				
	1	1		ļ		LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG
L		 	L			KT
1289	2639	Α	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT
1		l				DLREIGHGSFGAAYFARDVRTNEVVAIKKMS
	<u>L_</u> _	<u> </u>			<u> </u>	YSG
1290	2640	Α	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG
		1				TEATRPTAMSKSLKKKSHWTSKVHESVIGRN
]	j	J		j		PEGQLGFELKGGAENGQFPYLGEVKPGKVAY
	1	l				ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI
	1	}				KHCKDPLRLKCVKOGESSGLLSVLPGGGTAR
						GAGO
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLOPOASOOR
1271	2011	1 ^	10,10	120	371	RRRLPGRRQVQVQEGGGSGLRAWVLAMASV
1	ļ	l				LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK
1	1	l				
						VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE
<u></u>		<u> </u>				VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	Α	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR
}	ł	1				VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG
	ł	Ì				SWESWCCCCLIPADRPWDRGQHWQLEMADT
1	İ					RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM
}	1	ŀ	1			LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW
j	}	}	1			DAWSSLGDMTKEEAMIAYVEEMKKIIETMP
	ĺ	İ				MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD
	<u> </u>			<u> </u>		

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						LGNVLTSTPNAKTVNGKAESSDSGAESEEEE AC
1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL CGRGIRDSARMCSTCACVEYYGKALEVLVK GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV CTDVMARGIDIPEVNWVLQYDPPSNASAFVH
-						RCGRTARIGHGGSALVFLLPMEESYINFLAIN QKCPLQEMKPQRNTADLLPKLKSMALADRA VFEKGMKAFVSYVQAYAKHECNLIFRLKDL DFASLARGFALLRMPKMPELRGKQFPDFVPV DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN EGRRKFIKNKAWSKQKAKKK
1294	2644	A	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF
		,				GVLFALCFSCLLAHASNLVKLVRGCVSFSWT TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFLMALTFFVSKAT FCGPCENWKQHGRLIFITVLFSIIIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ
1295	2645	A	10133	376	518	HSFQVENQELSRDKWKVLLNSDFLSHSGA RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV
1296	2646	Α	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT
1299	2649	A	10161	1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV
1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV
1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alanine C-Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
		1100				
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ì		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
İ	l '			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	į		ľ	peptide	•	/=possible nucleotide deletion, \=possible
1	Ì			sequence		nucleotide insertion
	<u> </u>		ļ	sequence		KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
1						KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
1	i					DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
1	ļ					TERKSKHRNERKLSVLGKDGKPVSEYIIKTDE
1						NVRKENNKKERRLSAEKTKAEHKSRRSSDSK
1	İ					IOKDSLGSKOHGITLORRSESYSEDKCDMDST
						NMDSNLKPEEVVHKEKRRTKSLLEEKLVLKS
1						
						KSKTQGKQVKVVETELQEGATKQATTPKPD
1						KEKNTEENDSEKQRKSKVEDKPFEETGVEPV
						LETASSSAHSTQKDSSHRAKLPLAKEKYKSD
1	[KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD
1						ENKSDDKDGKEVDSSHEKARGNSSLMEKKL
1	İ					SRRLCENRRGSLSQEMAKGEEKLAANTLSTP
1						SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE
						NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
						ATVOKDELRTCTADSKATAPAYKPGRGTGV
1						l •
1						NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE
1						KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
1						QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
)						LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
						STSPADHSALPNQSLTVRESEVLKTSDSKEGG
1						EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
	i					GKVIMPLGSKLTGVIVENENITKEGGLVDMA
1						KKENDLNAEPNLKQTIKATVENGKKDGIAVD
}						HVVGLNTEKYAETVKLKHKRSPGKVKDISID
1	1					l l
1						VERRNENSEVDTSAGSGSAPSVLHQRNGQTE
1						DVATGPRRAEKTSVATSTEGKDKDVTLSPVK
1						AGPATTTSSETRQSEVALPCTSIEADEGLIIGT
i .	!					HSRNNPLHVGAEASECTVFAAAEEGGAVVTE
						GFAESETFLTSTKEGESGECAVAESEDRAADL
1						LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE
						KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG
1		·			l	TEIRAGSISSEEVDGSOGNMMRMGPKKETEG
1						TVTCTGAEGRSDNFVICSVTGAGPREERMVT
1 1			1			GAGVVLGDNDAPPGTSASQEGDGSVNDGTE
						GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
] [SESEENGESAMDSTVAKEGTNVPLVAAGPCD
1 1		(· .	Í		
						DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH
1						ASTCTGLGEESEGVLICESAEGDSQIGTVVEH
]			VEAEAGAAIMNANENNVDSMSGTEKGSKDT
				•		DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE
į i		- 1				GPMTSAASDQSDSQLEKVEDTTISTGLVGGS
1 1		i	1			YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE
1		ļ				NEECDGLMATTASGDITNONSLAGGKNOGK
1 1]				VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE
) {		.				ENMEGTRVTTEEFEAPMPSAVSGDDSOLTAS
[]		- 1				
1 1	1	1			ľ	RSEEKDECAMISTSIGEEFELPISSATTIKCAES
			}			LQPVAAAVEERATGPVLISTADFEGPMPSAPP
j		}				EAESPLASTSKEEKDECALISTSIAEECEASVS
; I						GVVVESENERAGTVMEEKDGSGIISTSSVEDC
			1		·	EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS
		1				TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA
1						AIISTSTAECMPISASIDRHEENQLTADNPEGN
						GDLSATEVSKHKVPMPSLIAENNCRCPGPVR
		ا.			ĺ	
		ľ			ļ	GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH
]]		ļ			ļ .	PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL
,)		1	J		ļ	HLINAEEKNVLLNSLQKEDKSPETGTAGGSST
1 [j		į	ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK
j ļ		ļ	İ	l		DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA
į i	i	İ	1	1	i	EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP
		ı				

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ł	l	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	I	ļ.	[peptide		/=possible nucleotide deletion, \=possible
	 	 	 	sequence		nucleotide insertion
1		1	\			SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS
ļ]	J]		SEENVCDIGNEESPLNVLGGLKLKANLKMEA
			1	l		YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE
1	ļ]		PLLVNESLNVENSGFRTNEEIHSESYNKGEISS
1	1	1				GRKDNAEAISGHSVEADPKEVEEEERHMPKR
İ		1		1		KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC
		1	1	1	İ	PETEPHATKEENSRDLEELPKTSSETNSTTSRV
L		<u></u>	L		<u> </u>	MEEKDEYSSSETTGEKPEQNDDDTIKSQE
1302	2652	A	10167	321	842	EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH
		!	[FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG
		1	1			NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE
1	1	!	1			DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP
		1				EARETKCYVRSSVGCVEPLTTQAEVTENLDR
1303	2653	A	10171	206	429	KNSQQVFKLLKKK NMILLKKRRLLINSLGEGTINGLLDELLETNV
1303	2005	^	101/1	200	429	LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA
	l	ł	!			RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAONGF
1307	2007	ļ ^`	10.07	1 770	1324	YHEAVVLFTQALKLNPODHRLFGNRSFCHER
1		1				LGQPAWALADAQVALTLRPGWPRGLFRLGK
1	1	ļ	1	1		ALMGLQRFREAAAVFQETLRGGSQPDAAREL
		[l		RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA
		L				ELAPSGLPSLRCPRSTALRSPGLSPLLH
1305	2655	Α	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ
}	j	j	ļ]		SDFLTPPVGGAPWAVATTVVMYPPPPPPHR
]]	1	}			DFISVTLSFGESYDNSKSWRRRSCWRKWKQL
		İ	1			SRLQRNMILFLLAFLLFCGLLFYINLADHWKG
1306	2656	A	10195	1	410	IRNTCT IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY
1200	2030	^	10133	'	110	NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI
]	l]			GIDDEDDSTFTTTVDQKTFHFQARDADEREK
	[1	ł].		WIHALEETILRHTLQLQVRVFTWFPDSSLVGA
]	<u> </u>	١.			FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
	1	l	ł	i		DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS
L	L	<u> </u>		L		SPDQQNYTKSR
1308	2658	Α	10214	2	453	ECGGIRQPGPPPPALASAPAATMNRVGGSPS
1	ł	ł			!	AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP
1	1	l				SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL
ł	ł	l	1	l		DNITQVMSLHTQYLESFLRSQFYMLRMDGPL
1309	2659	A	10233	45	421	PLPYRHYIAIMAAARHQCSYLINM RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
1309	2039	^	10233	43	421	RALESRTFQGSERSRWGPPLESTKENVQCGH
	1	l	l			RPAFPNSSWLPFHERLQVQNGECPWQVSIQM
		l				SRKHLCGGSILHWWWVLTAAHCFRRTLLDM
	1	l	l			AV
1310	2660	A	10241	243	442	AFOLFNAKCESAFLSKRNPLORNWTVLYRRK
		l	1			HKKGQSAEIQKKRTRRAFKFQRAITGASLADI
	1	ĺ		1		MAK
1311	2661	A	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ
		l		· ·		VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ
	J		j]		LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
		l				QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI
	<u> </u>		!			DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML
	ſ	[ĺ			QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI
1312	2662		10270		660	KAG CTCCDCCCDCACTBAILLYTGEVECAEVDVIEVD
1312	2662	Α	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP SMTILDKKDGEQAKALFEKVRKFRAHVEDSD
L	<u> </u>	<u> </u>	L		L	OWITH THE PROPERTY OF THE PROP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	<u> </u>	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	1	09/496 914	correspondi ng to first	to last amino acid residue	M.—Methionine, N.—Asparagine, P.—Proline, O.—Glutamine, R.—Arginine, S.—Serine,
dence	1	f		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}	ļ	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	<u> </u>	!		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	}					LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF
		j				EHVCKPKVEHLIGYEVFECTHNMAYMLKKL LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV
	ł	1				REESSFSDIPDVKNDFAFLLHMVDQYDQLYS
			!			KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	A	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR
İ	İ	i				VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL
			!			LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE
	İ	1	!		(EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD
1		1				GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL
i i	İ	1				PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW
		i				GRVKNFLMFFSFSRSOWDMOSTFDNTIKEHF
	Ì			ŀ	ļ	PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL
						LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	Α	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL
			1			VKKDKSEKELKALCIDQLDVFLQKETQIFVEK
		1	1			LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR
	ì	ł	İ I			ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD
	}	1				RYNRRRGRSRSYSRSRSRSWSKERLRERDRD
		1			[RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN
1		}				NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG
	İ					NNTTESWSEFHEDQVDHNSYVRPPMPKKRC
		-			:	RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN
	·	1			}	LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPPV NLRPPVPPPGPLPPSLPPVTGPPPPLPPLQPSG
	1	ļ				MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT
	Ì	l				ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR
		ŀ				AKLG
1315	2665	Α	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA
	ĺ	1	İ			GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI
}	j	ļ			J	GTPSPCGSAQAAAAAAAEEATEKIPALRPALL WALLALWLCCATPAHALOCRDGYEPCVNEG
						MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE
1	1	l]			KNRCQNGGTCVAQAMLGKATCRCASGFTGE
1						DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE
]				1	CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS
1	1	ł			}	GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT
1316	2666	A	10294	118	572	LWVMVSGATSTSTSTL SLSMESNHKSGDGLSGTQKEAALRALVQRTG
1310	2000	[^	10294	110	3/2	YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI
	ľ	l				GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF
}])]		ļ	NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR
						NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	Α	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ
[]	1			Ì	DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP
1	l	l				LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES
]			RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV
1	ĺ]			PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP
ľ	1	ĺ				KQENEGGEKAALESQVRELLAEAKAKDSEIN
}]	}]]		Ì	RLRSELKKYKEKRTLNAEGTDALGPNVDGTS
		1				VSPGDTEPMIRALEEKNKNFQKELSDLEEENR
	!	1				VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ
1.	İ]				ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS
1		1				NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM
1	}	ł			}	EENHHSTAEELQATLQELSDQQQMVQELTAE
<u> </u>	L			L	L	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
donot		ĺ	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			l	peptide	acquence	/=possible nucleotide deletion, \=possible
Ì	ļ			sequence		nucleotide insertion
J	 	 		Sequence	<u> </u>	NEKLVDEKTILETSFHQHRERAEQLSQENEKL
}]			ļ	MNLLOERVKNEEPTTOEGKIELEOKCTGILE
	l		!			QGRFEREKLLNIQQQLTCSLRKVEEENQGAL
		1				
i i		•				EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1210	2660		10202	222	879	EECRVTLEGLKMENGSLKSHLQG
1318	2668	A	10303	333	879	GECFIMAAVVQQNDLVFEFASNVMEDERQL
1	1					GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
İ						NDMITESSLDVAEEEIIDDDDDDDITLTVEASCH
1		İ	ĺ			DGDETIETIEAAEALLNMDSPGPMLDEKRINN
İ					j	NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
L						QVQEKYADSPGASSPEQPKRKKK
1319	2669	Α	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLLIJATALSL
	Ì		ļ			LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR
}						LPRTLAGLLAGGALGLAGALMQTLTRNPLAD
	İ	ļ	· '			PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
	ł	1	ļ			MAFAGALVASLIVAFTGSQGGGQLSPVRLTL
1	}	l				AGVXL
1320	2670	Α	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
j	}	ļ				AVVDIQSDKAANVAQEINAEYGESMAYGFG
		1	1			ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
1	ļ		ļ			AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
1	1	1	·			FSRLMIRDGIQGRIIQINSKSDE
1321	2671	Α	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
		*-		_		AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
	İ					ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
	ļ					YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
}						MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
17-2	1 20/2	l	10333		123	ERKMRAHQVLTFLLLFVITSGASENASTSRGC
l	Ì		(GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
1				'		LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL
}		Į				LGTLG
1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
1323	20/3	^	10334	32	420	QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
		ł				YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI
[[LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	A	10336	1	932	L
1324	2074	1	10220	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
		,				NSVTHHEVKCQGKPLAGIYRKREEKRNAGN
		ļ			ı	AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA
		'				AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
						PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
		į į				LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
					į l	VIATKQFSRGDFVVEYHGDLIEITDAKKREAL
						YAQDPSTGCYMYYFQYLSKTYCVDATRETN
					,	RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
						RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	Α	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
		[[RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
						PMRKSYRGDREAFEETHLTSLDPVKQFAAWF
					,	EEAVQCPDIGEANAMCLATCTRDGKPSARML
						LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
						VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
]						RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
				1		QLYQDQEVPKPKSWGGYVLYPQVMEFWQG
						QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
					İ	DWLYERLAP
1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
]	· -			-		LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT
1						LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA
	L,		اا			

SEO ID	SEQ ID	Met	SEQ	Dandistad	Predicted end	I Amino onid an anno (AmAlonius Co-Cartaina
			, ,	Predicted		Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ì	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1]	j	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	1	peptide	· ·	/=possible nucleotide deletion, \=possible
1	i	İ	ľ	sequence	ĺ	nucleotide insertion
	 		 			HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD
}		ļ	ł	İ	}	EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY
	1		ļ	Ì		LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP
	İ	İ	ļ	Ì	<u> </u>	NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA
		j			į	,
	ł]	l	i	1	ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ
}	l	ļ	i		(ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD
i	Ì	l	1		İ	MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
L		<u> </u>	L			LDKLGYSQKMRFTY
1327	2677	Α	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
			1			RYHRLSDMSMLAERRRKQKWAVDPQNTAW
1	}					SNDDSKFGQRMLEKMGWSKGKGLGAQEQG
1	1	[ATDHIKVOVKNNHLGLGATINNEDNWIAHO
	l	1			ļ	DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
1	1	1	1			LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL
1	1	1				
	1	ļ				DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF
1	ĺ	ĺ			[TIQEYFAKRMAALKNKPQVPVPGSDISETQVE
	ŀ	ļ				RKRGKKRNKEATGKDVESYLQPKAKRHTEG
1	l	1	1		ļ	KPERAEAQERVAKKKSAPAEEQLRGPCWDQ
						SSKASAQDAGDHVQPA
1328	2678	Α	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI
1		ĺ	Į.			CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
1	[l	İ			HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	A	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL
	1			_	100	SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN
	ļ		j			LSFADICVTSTTIPKMLMNIQTQNKVITYIACL
	1	İ	1			MOMYFFILFAGFENFLLSVMAYDRFVAICHP
1	ł					LHYMVIMNPHLCGLLVLASWTMSALYSLLQI
	l	ì	l			
	l	l	1			LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF
1		j)			LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH
	i					AISSAQGKYKAFSTCASHLSVVSLFYGAILGV
		İ				YLSSAATRNSHSSATASVMYTVVTPMLNPFI
			Ĺ			YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	Α	10352	34	2573	IPFLKSCCCCCLFDFPPPPLDQVQEEECEVERV
1	ł i	1	1			TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE
		1				TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL
]]	1	}			FYTERAHVRTLKVLDQVFYQRVSREGILSPSE
1						LRKIFSNLEDILOLHIGLNEOMKAVRKRNETS
		1	1			VIDOIGEDLLTWFSGPGEEKLKHAAATFCSNO
1		l				PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR
		1				LOLKDIIPTOMORLTKYPLLLDNIATYTEWPT
1)]			EREKVKKAADHCROILNYVNOAVKEAENKO
]			, , , , , , , , , , , , , , , , , , , ,
	į	1				RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK
		ĺ	(RKMIHEGPLVWKVNRDKTIDLYTLLEDILV
						LLQKQDDRLVLRCHSKILASTADSKHTFSPVI
1		1		,		KLSTVLVRQVATDNKALFVISMSDNGAQIYE
Į l		1				LVAQTVSEKTVWQDLICRMAASVKEQSTKPI
		1]			PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL
]		1	,	1		QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD
\			i			LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
1 1			'			HLPVSEERWALDALRNLGLLKOLLVOOLGLT
1						EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE
1			ļ ļ			NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE
						SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
1						MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
		l				
1			(!	GDGAVNKEEKDVNLRISGNYLILDGYDPVQE
1		ľ				SSTDEEVASSLTLQPMTGIPAVESTHQQHSP
1						QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI
			ļ i			QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE
L						SYTILCQRLAGSALTDKHSDKS

SEQ ID NO: of nucl- nucleotide seq- uence uence uence uence uence
nucleotide sequence nucleotide corresponding nucleotide sequence nucleotide sequence nucleotide sequence nucleotide sequence nucleotide sequence nucleotide insertion nucleotide
eotide sequence uence
sequence Sequence
uence 914 ng to first amino acid residue of peptide sequence Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion nucleotide insertion AVEFAEGALTMAPWPELGDAQPNPDKY AAGQQPTAPDKSKETNKTDNTEAPTKY SYSTATLIDEPTEVDDPWNLPTLQDSGIK RDTKGKILCFFQGIGRLILLLGFLYFFVCS SSAFQLVGGKMAGQFFSNSSIMSNPLLGI VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR, GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVC KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTIGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVLLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVVK FDNITISREAQGEVFASDSKTECTAL
amino acid residue of peptide sequence peptide sequence s
peptide sequence
Sequence nucleotide insertion
1331 2681 A 10353 I 2100 AVEFAEGALTMAPWPELGDAQPNPDKY AAGQQPTAPDKSKETNKTDNTEAPVTKI SYSTATLIDEPTEVDDPWNLPTLQDSGIK RDTKGKILCFFQGIGRLILLLGFLYFFVCS SSAFQLVGGKMAGQFFSNSSIMSNPLLGV VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR. GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILIS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLD DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
AAGQQPTAPDKSKETNKTDNTEAPVTKI SYSTATLIDEPTEVDDPWNLPTLQDSGIK RDTKGKILCFFQGIGRLILLLGFLYFFVCS SSAFQLVGGKMAGQFFSNSSIMSNPLLGI VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRRA GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVC KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
SYSTATLIDEPTEVDDPWNLPTLQDSGIK RDTKGKILCFFQGIGRLILLLGFLYFFVCS SSAFQLVGGKMAGQFFSNSSIMSNPLLGI VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR. GATVHDFFNWLSVLVLLPVEVATHYLEI IVESPHFKNGEDAPDLLKVITKPFTKLIVC KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL. ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
RDTKGKILCFFQGIGRLILLLGFLYFFVCS SSAFQLVGGKMAGQFFSNSSIMSNPLLGI VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR. GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIIVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
SSAFQLVGGKMAGQFFSNSSIMSNPLLGI VLVTVLVQSSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR, GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL, ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIIIVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
VLVTVLVQSSTSTSIVVSMVSSSLLTVR IIMGANIGTSITNTIVALMQVGDRSEFRR, GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILL LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL, ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVQVVVFIIIVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
IIMGANIGTSITNTIVALMQVGDRSEFRRAGATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAILL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKAFDNITISREAQGEVPASDSKTECTAL
GATVHDFFNWLSVLVLLPVEVATHYLEI IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL, ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVI LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
IVESFHFKNGEDAPDLLKVITKPFTKLIVG KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL. ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
KVISQIAMNDEKAKNKSLVKIWCKTFTN INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL. ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF, LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
INVTVPSTANCTSPSLCWTDGIQNWTMK YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
YKENIAKCQHIFVNFHLPDLAVGTILLILS LCGCLIMIVKILGSVLKGQVATVIKKTIN' FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
LCGCLIMIVKILGSVLKGQVATVIKKTIN FPFAWLTGYLAILVGAGMTFIVQSSSVFI TPLIGIGVITIERAYPLTLGSNIGTTTTAIL, ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
FPFAWLTGYLAILVGAGMTFIVQSSSVFT TPLIGIGVITIERAYPLTLGSNIGTTTTAIL, ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
TPLIGIGVITIERAYPLTLGSNIGTTTTAIL ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
ASPGNALRSSLQIALCHFFFNISGILLWYF RLPIRMAKGLGNISAKYRWFAVFYLIIFF LTVFGLSLAGWRVLVGVGVPVVFIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
LTVFGLSLAGWRVLVGVGVPVVFIIIILVL LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
LLQSRCPRVLPKKLQNWNFLPLWMRSLI DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
DAVVSKFTGCFQMRCCCCCRVCCRACC GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
GCPKCCRCSKCCEDLEEAQEGQDVPVKA FDNITISREAQGEVPASDSKTECTAL
FDNITISREAQGEVPASDSKTECTAL
1332 2682 A 10354 30 1377 SQQGQPHRQGPPSLLTAPHSLDLPALPP
GSQGKLRRVLVPMSVKPSWGPGPSEGV
TSDLGEIHNWTELLDLFNHTLSECHVELS
KRVVLFALYLAMFVVGLVENLLVICVN\ SGRAGLMNLYILNMAIADLGIVLSLPVW
VTLDYTWLWGSFSCRFTHYFYFVNMYS
LVCLSVDRYVTLTSASPSWQRYQHRVRI
CAGIWVLSAIIPLPEVVHIQLVEGPEPMCI
APFETYSTWALAVALSTTILGFLLPFPLLT
VLTACRLROPGOPKSRRHCLLLCAYVAV
MCWLPYHVTLLLLTLHGTHISLHCHLVH
FFYDVIDCFSMLHCVINPILYNFLSPHFRO
NAVVHYLPKDQTKAGTCASSSSCSTQHS
KGDSQPAAAAPHPEPSLSFQAHHLLPNTS
TQPLTPS
1333 2683 A 10358 2 884 AAGAGADGREPASERASRAEPPAVAMG
LMGTAEDFADQFLRVTKQYLPHVARLCI
FLEDGIRMWFQWSEQRDYIDTTWNCGY
SSFVFLNLLGQLTGCVLVLSRNFVQYACI
GIIALQTIAYSILWDLKFLMRNLALGGGL
LAESRSEGKSMFAGVPTMRESSPKQYMQ RVLLVLMFMTLLHFDASFFSIVQNIVGTA
LVAIGFKTKLAALTLVVWLFAINVYFNA
IPVYKPMHDFLKYDFFOTMSVIGGLLLV
GPGGVSMDEKKKEW
1334 2684 A 10367 59 1562 QAWSLQVALSPFFFPASPSNSFAAAVPQL
ELPLPHVPGQESAKRRSARRFLIMSELTK
ELVWGTKSSPGLSDTIFCRWTQGFVFSES
ALEQFEGGPCAVIAPVQAFLLKKLLFSSE
WRDCSQEEQKELLCHTLCDILESACCDH
YCLVSWLRGKTTEETASISGSPAESSCQV
SALAVEELGFERFHALIQKRSFRSLPELKI
LDQYSMWGNKFGVLLFLYSVLLTKGIEN
EIEDASEPLIDPVYGHGSQSLINLLLTGHA
VWDGDRECSGMKLLGIHEQAAVGFLTLI
LRYCKVGSYLKISKIPYLDCLASETHLTV
KDMALVAPEAPSEQARRVFQTYDPEDNO
DSLLEDVMKALDLVSDPEYINLMKNKLD

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LGIILI.GPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
1335	2685	A	10375	82	2929	TRTKRRLGREKAMASPPRGWGCGELLLPFML LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDNDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISEIFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNANGKRDGKAPAGGNGN KKKSGKKEKK
1336	2686	A	10379	1	557	RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK LAKLQAQVRIGGKGTARRKKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380		1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

OPO TO	L CEC TE	1 14-4	CEC	D-divi	I Book of the control	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l .	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
)	1	peptide	; ·	/=possible nucleotide deletion, \=possible
				sequence	ŀ	nucleotide insertion
						QPN
1339	2689	Α	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG
J				i		KCVICDSYVRPCTLVRICDECNYGSYQGRCVI
			}			CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL
	ļ		<u> </u>			GSSKTDLFYERKKYGFKKR
1340	2690	Α	10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK
		İ	1			RPSWMVDNKRMRTASNFQWLLSTFILLYLM
		1		j	ļ	NQVNSQKKGAPHDLKCVTNNLQVWNCSWK
		ŀ	1			APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP
			ļ	}		ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL
		1	1	ĺ		IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR
		ì	1	Ì		SNVIWEIKVLRKESMELVKLVTHNTTLNGKD
			ŀ			TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSBYKNISWIBDSOTKVEBODKVII
	ļ	}				GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE
			l			NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF
						AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA
			ļ			LVGPRATSYTLVÉSFSGKYVRLKRAEAPTNES
		l	[ĺ		YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL
			1			VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP
		l	İ	}		GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE
		}	ł			NSSYLVALDKLNPYTLYTFRIRCSTETFWKW
	1	İ				SKWSNKKQHLTTEASPSKGPDTWREWSSDG
			ļ			KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ
		J	ļ		j	SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS
		1	· ·			SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW
			Ì			HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV
						PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY
		-	1			QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV
		ļ	ļ .			KWEDIPVEELRGFLRGYLFYFGKGERDTSKM
	ľ	1				RVLESGRSDIKVKNITDISQKTLRIADLQGKTS
						YHLVLRAYTDGGVGPEKSMYVVTKENSVGL
		l			,	IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP
						CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN
		[RSDAKPENHVVESYCPPIIEEEIPNPAADETGG
	1					TAQVIYIDVOSMYOPQAKPEEEQENDPVGGA
i	1	l	1			GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP
						QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI
	1	l				NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP
<u> </u>	ļ	})			ND
1341	2691	Α	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP
l	1	1				DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL
	{	ĺ	[IKLRNLCLDWLQPETRTKEEHELLVLEQYLTII
		ĺ	1			PEKLKPWVRAKKPENCEKLVTLLENYKEMY
						QPEGESLHGVLVVSAGLRCPLGLSASTLLTW
	I	ľ				SGLDNSLSWAAVGMSCVLWDIELHHDFLGV
		Ì				ATKSVSTHAQGDAAQGLGGTTVRMWARDSN
]		j			LATGVLLDDNNSDVTSDDDMTRNRRESSPPH
						SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR
			[NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS
						RAYESRSQDAESYQNVVDLAEDRKPHNTIQD
		{	l ·			NMENYRKLLSLGVQLAEDDGHSHMTQGHSS
						RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED
			Į l			ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ
		l	j .			RFPRMSDDNWKDISLNKRESVIQQRVYEGNA
						FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK
						GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS
		ł				SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS
	L	L				EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

NO. of mucle olde on the control of	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
peptide colide sequence unner contemporaries and process of the corresponding of the sequence				. ~			
uence 09/496 Oorrespondi on glo finst amino acid residue of peptide residue of peptide sequence O-Glusmine, R-Arginine, S-S-Serine, of peptide sequence O-Glusmine, R-Arginine, S-S-Serine, of peptide sequence O-Glusmine, R-Arginine, S-Serine, of peptide sequence O-Glusmine, R-Arginine, S-Serine, of peptide sequence O-Glusmine, R-Arginine, S-Siop codon, P-possible nucleotide deletion, Pspossible nu	nucl-	peptide	}		nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
Part Barrio Bar	eotide	seq-		USSN	location	corresponding	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per control of the per control of the period of period		uence		I			
residue of peptide sequence y=Tyrosine, x=Unknown, *=Siop codon, /posible nucleotide delicin, y=posible nucleotide insertion SQVGGKRPECKDCGETFINKSALALAFIRKIHA ROYLVECKNOGECEAPMPSITISELQKIYOK DKYYECKVCETFILMSSALEHOKHRIGDDKD NERHIBERERERGETFRFSALNEFQKMYG KEKMYECKVCEGTFILMSSALEHOKHRIGDDKD NERHIBERERERGETFRFSALNEFQKMYG KEKMYECKVCGETFILMSSALEHOKHRIGDDKD NERHIBERERERGETFRFSALNEFQKMYG KEKMYECKVCGETFILMSSALEHOKHRIGDDKD NERHIBERERERGETFRFSALNEFQKMYG KEKMYECKVCGETFILMSSALEHOKHRIGHT STENDEDEKA STENDERAPMQSSELSHOKHRISKNIHFOR OYEKSVIHISGPTESOSINTRIPLISDDEEKA STENDERAPMQSSELSHOKHRISKNIHFOR OYEKSVIHISGPTESOSINTRIPLISDDEEKA STENDERAPMQSSELSHOKHRISKNIHFOR STENDERAPMQSSELSHOKHRISKNIHFOR STENDERAPMSSELSHOKHRISKSILSHOKHRISKSILSHOKHRISKSUCH STENDERAPMSSELSHOKHRISKSILSHOKHRISKSILSHOKHRISKSILSHOKHRISKSUCH STENDERAPMSSELSHOKHRISKSILSHOKHRISKSILSHOKHRISKSILSHI STENDERAPMSSELSHOKHRISKSILSHOKHRISKSILSHI STENDERAPMSSUCHISTOR	uence]	į	914			
peptide sequence Popsible incertain Popsible							
sequence nucleotide insertion	1	ł	l			sequence	, , , , , , , , , , , , , , , , , , , ,
SOYGGKRFECKDCGETFINSSALEHRKIHA RGYLVECKNOECELAFMPSITESLCKIYOK DKFYECKVCKETFIHSSALIEHOKIHGDDKD NEREHERERERREGTFRSPANLINEFOKMYG KEKMYECKVCGETFIHSSALIEHOKIHGDDKD NEREHERERERREGTFRSPANLINEFOKMYG KEKMYECKVCGETFIGSSLKEHOKIHTRGN PENKGKVCEETFINSSALISHOKHTRGHT DFDCKTORDAMOGSELSEHOKHRSKNILFEGR GYEKSVHISGPTESOKSHTITRILESDEEKA FTISSNPYENOKIPTKENVYEAKSYERSVHISL ASVEAQKSHISVAGPSKYMAESTICSBDENA HORVEAGGINTSGGRYSRSVHISLVASKPPSS HIGHLYSNEKGESSIYSIDINDKOKIPAR ENPICEGGSKNRIYEDSVIGSVFRAKPOKSVP GGSGSEKLOGESTSYSSNVHSLVASKPPSS HIGHLYSNEKGESSIYSIDINDKOKIPAR ENPICEGGSKNRIYEDSVIGSVFRAKPOKSVP GGSGEKLOGESTSYSSNVHSLVARARKK YÜBHRSNETSVHISLPFGGOTFRRGMLYECO ECGECHAISSDLTEHOKIDHERPSCSINVE WSVIRSLAPTDOTISVAGEOVAKEQARNKCK VÜBHRSNETSVHISLPFGGOTFRRGMLYECO ETGECHAISSDLTEHOKIDHERPSCSINVE WSVIRSLAPTDOTISVAGEOVAKEQARNKCK VÜBHRSNETSVHISLPFGGOTFRRGMLYECO ENPICEAHSSDLTEHOKIDHERPSCSINVE WSVIRSLAPTDOTISVAGEOVAKEQARNKCK VÜBHRSNETSVHISLPFGGOTFRRGMLYECO ENPICEAHSSDLTEHOKIDHERPSCSINVE WSVIRSLAPTDOTISVAGEOVAKEQARNKCK VÜBHRSNETSVHISLPFGGOTFRRGMLYECO ENPICEAHSSDLTEHOKIDHOKIPACH KOCCO ENPICEAHSSOLTEHOKIPACH KOCCO ENPICEAHSSOLTEHOKIPACH KOCCO ENPICEAHSSOLTEHOKIPACH KOCCO ENPICEAHSSOLTEHOKIPACH ENPICEASSALVITHON WILKIGGLINVERALERPSCALERPSCALERPS GAMERPTECKALERPSCALERPSCALERPS GAMERPTECKALERPSCALERPSCALERPS GAMERPTECKALERPSCA	1			ŀ			
RGYLVECKNOECEEAMPSPITTSELQKIYOK DKPYECKVCKETH-HEDDKD NEREHERERERGETFRPSPALNEFQKMYG KEKMYECK VCGETTH-HESSILEHOKHITGDNKD NEREHERERERGETFRPSPALNEFQKMYG KEKMYECK VCGETTH-HSSILEHOKHITGNK-HEDGK DFTDGRAPMQSSELSHQKHERKNI-HEGR GYEKSVIHSGPTESOKSHTITTR-LESDEDEKA FITSSPYPENOKLIPTKEN-YEAKSYERSVIHSL ASVEAQKSHSVAGPSKEKWARESTICSEDADN HQRVFAGGOTTSGERSYSSWIHSL-VASKSPPBS HNONELVESNEKGESSIN'SDLNDKRQKUPAR ENPCEGGSKNRNYEDSVIQSVPRAKPQKSVP GGSGERKNOGFSVPSSNVREVQKARAKKK YTEHRSNETSVHISL-PTGEQTTFRROML-NECQ EGGECT-AISSDLETHQKHIDENERYGSSNV- WSVIRSLAPTDPQTSY AQEQYAKEQARNKCK DPRQFFATSEDLNTNGKIPOKESTGESGSGNV- WSVIRSLAPTDPQTSY AQEQYAKEQARNKCK DPRQFFATSEDLNTNGKIPOKESTGESGGE- NTDGEETHSETHQETTEDPVIQGSDMEDPQ KDPPDOKINEEDEGGFVDLTD-ITOHQK-H SRKCLVDSREYTHSVHTHSISEVQRDYTGEQ LYECPKCGSFRIHSST-HERQRHEQDLYSM KGCDDGFHALLPMKPRRNEAAERNPALAGSA RICLLGGQFHISSALHERHRH-RIEDDLEQS QMAEEAINFGLALTFQRSQTEERIFFCAVCG ESFVNRAELADHVTVHKNEPYEVGSVSYTHTS FITSPLKGAJPFYECKDCGKSFIHSTVLTKHKE LHLEEGEBEAAAAAAAQVEVANVHVPQ VVLRIQGLNVBAAEPSVEAABPSVEAABPEVAAAPFVE EAAEPNGAEAGPDGEAGPGGAGPOKGAB QPNGDADEPDGAGIEDPERRAEPEGKAEPEG GDADEPPDGAGIEDPERRAEPEGKAEPEG GDADEPPDGAGIEDPEGRAEPEGGRAEPEG GDADEPPDGVGIEDPEGGLOGIGOVEPYYDC HECTETTISTSTAPSEHLKTHASMIFFPANAFG ECSGYIERASTSTGGAAAQAQEKYKTOCHOG LINDHISLARRIQNTHTG GFRSSSSDNRYFLERERGGLSSAAVQTRIGNSA ASPANLAVVAVAVVVVVVVAVVAVAVANANYHVP PGSFEVYRLNYTVOODGEHECREGALISLLQ AHGEGROVFDAYSTGANTALVTENDEVSSIFFV GLWALIQAKYSTIETDFGLOFFGINKRICKVGG TMEDVVLVRIYGNKTELLDDRDEVSSIFFV GLWALIQAKYSTIETDFLOCYNGHNENE PAGVSDDVVSLYHORDEVSSIFFV GLWALIQAKYSTIETDFLOCYNGHNENE PAGVSDDVVSLYHORDEVSSIFFV GLWALIQAKYSTIETDFLOCYNGHNENE PAGVSDDVVSLYHORDERVSSIFFV GLWALIQAKYSTIETDFLOCHNENE PAGVSDDVSNIYHPRELIGONGNENE PAGVSDDVSNIYHPRELIGONGNENE PAGVSDDVSNIYHPRELIGONGNENE PAGVSDDVSNIYHTSVBDPSOI LOGEMTWAKELISHI GSPVVLCHNDLLCKNII VYEKGODOVSTONTYSTAVANDAVOV WORKCDMEDBEVSTSRAVL GWALGARSSLAFFFW GLWALIQAKYSTIETDFGRUDHCRAE LINDHISSENEN DOWSTONTYSWDNAQ WILVGNKCDMEDBEVSTSTRAGHLIGGGLGFF FITSSKAKKININVKQTFELUDILICSKOCK TOPATHALKYPE PAGTSTAVANDSTTASTAVYTVOODEFEKGEGLIPKH MYPTITATAYPROGRIFTE PAGTSTAVANDSTYTASTAV		Ļ	<u> </u>		sequence	ļ	
DIKFYECKYCESTFLHSSALIEH()KIHIGDIND NEREHBERERREGFERSPENDER/SMYG KEKMYECKYCETTFLOSSLIKEH()KIHITGEN NEREHBERERREGFERSPENDELKA PROKINTER SENDEN KEKMYECKYCETTEN SESSLEH()KIHITGEN PERMKEK VCCETTEN SESSLEH()KIHITGEN GYEKSYHIGEPTES SCHITTRIVES DELEKA PERMEKEN SCHITTRIVES DELEKA FITISNPYENQKIPTKEN YEAKSYERSVIHSL ASYEAQKSHIS WAGESRIKYMAESTIQSTDAIN HQRVRAGONTSEGRESYSRAVHSL VASSPERS HRONEL VESNEGESSIN SUNDKRÇKIPAR ENCEGESKRIKYMES VIGNOK KIPTAR ENCEGESKRIKYMES VIGNOK KIPTAR ENCEGESKRIKYMES VIGNOK KIPTAR ENCEGESKRIKYMED SVIGNOK KIPTAR ENCEGESKRIKYMED SVIGNOK KIPTAR ENCEGESKRIKYMES VIGNOK KIPTAR ENCEGESKRIKYMED SVIGNOK KIPTAR ENCEGESKRIKYMED SVIGNOK KIPTAR ENCEGESKRIKYMED SVIGNOK KIPTAR ENCEGES VIGNOK KIPTAR ENCEGES AND SET STAND		l	}	İ]	J	
NEREHERERERGETFRSPALNEFOKHTYGG KEKMYECK VCGETFLISSISLEHOKHTRYGG REPNIGG VCETTFLIGGSILKEROKTYNIKEL DFTDGDAPAMOSESLEHOKHTRINEDOE GYEKSVHISOPFTESOKSHTITRI ESDEDEKA FITSSNYENOKUTKENYVEAKSYERSVHISL ASVEAOKSHSVAOPSKPKYMAESTIOSFDAIN HQXVRAGGNTSEGREYSRSVHISLASSPRSS HNGNELVESNEKGESSIYSDLINDKROKDRAW ENPECGGSKRINVEDSVIQSVFRAKPOKSVP GEGSGEFKKDGEFSVPSSIVIRTYOKASKPRS HNGNELVESNEKGESSIYSDLINDKROKDRAW ENPECGGSKRINVEDSVIQSVFRAKPOKSVP GEGSGEFKKDGEFSVPSSIVIRTYOKASKPRS HNGNELVESNEKGESSIYSDLINDKROKDRAW ENPECGGSKRINVEDSVIQSVFRAKPOKSVP GEGSGEFKKDGEFSVPSSIVIRTYOKARK YBEHRSNETSVHISLIFFGGFTEIBPRIGNILTEOK DFROMFATSELLINTINGKIYTQREKGESGGE NTDGETTISSETHGGFTEIBPPVIQGSDMEDDO KDDPDDKTYEEGDCGLGFVDLIDTHOKOKT SKRCLVDSRSYTHSVHITHSISETYQRDYTGGC LYECFKCGSSHTSSELFEHQRHEDQOLYSM KGCDDGFTALLFMKFRRNRAAERNPALAGE LYECFKCGSSHTSSELFEHQRHEDQOLYSM KGCDDGFTALLFMKFRRNRAAERNPALAGE LYECFKCGSSHTSSSLEFHQRHEDQOLYSM KGCDDGFTALLFMKFRRNRAAERNPALAGE QMAEBAIBGLALTETGRSGTEERLFECAVCC GESFVNPAELADHVTVHKNEPYEVGSSYTHTS FILTEPLKGAIFPYECKDCOKSFHISTVLTKHEE LHLEEEEDBAAAAAAAAACPEVVEAAPPEVEAAPPEV EAAAFPNGEABGPDGAAEPIGGAQPNGEAB QPNGDADEPDGAADFIGGAQPNGEAB QPNGDADEPDGAADFIGGAQPNOEAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAQPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPDGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGGAGPNOEBAB QPNGDADEPGGAADFIGAABAB QPNGDABABABABABABABABABABABABABABABABABABAB	1		1	1			
KERMYECK VCGETFLHSSSLKEHOKIHTRICK	}		}	<u> </u>			1
PFENNGKVCETTIPIGGSLKRRQKTYNIKELG DFTDGDAPMGSSELSEHQIKHRRNLFEGE GYEKSVHISQPFTESQKSHTTRRI ESDEDEK A FITSSNYENQKTYKENYPEAKSYERSVIHSI ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN HQRVRAGGONTSEGREYSRSVHISLVASKYPRS HNONELVESINGKGSIYISDLNSKQKIPAR ENNCEGGSKRRINTEDSVIQSVFRAKPQKSYP GEOSGEFKKDGESYYSSNYREYQKARAKKK YEHRISNETSVHISLIPFGEQTFRPRGMLYECQ ECGECFAHSDLTEHQKHIPDEKPGSSNYPE WSVIRSLAPTDPQITSVAQEQYAKEQARNKK THERSPETSVHISLIPFGEQTFRPRGMLYECQ ECGECFAHSDLTHOKQKIPDEKSGSGSQG NTDGETHISETHGGETEDPVIQGSDMEDDQ KDDPDDKIYECEGCGGFVLDITDHOKQVH SRRCLVDSRSYTHSVHITHSISEVQRDYTGGC LYECPKCGESFHISSTLEHQRHIPERDQL YSM KGCDDGFIALIPMKPRRNRAAERNPALAGSA IRCLLCOQGHISSTLEHQRHIPERDULEQS QMAEEAIPGLAATFORGTEERLFECAVCG ESFNYBAEADHVTVHKNEPYEVGSSYTHIS FILTELKGAIPFYECKDCCKSFHISTVLTKHIKE LHLEEBEEDBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAFPEVAAEPFUKAAEPEV EAAEPNGEAEGPDGAAEPFUKAAEPEVGAAEPE GDADEPDOVGIEDPEGGEDQEQVEEPYYNO HCCTETTISTSTAFSHILKTHASMIFPBANAFG ECGGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLAHQNHTHG ECGGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLAHQNHTHG FROM THE PROPERTY OF TH				İ			l , , , , , , , , , , , , , , , , , , ,
DFTOGRDAPMQSSELSEHQKIHSRKNLTEGDERA GYSKSVHISOPTIESOSQEHTTRIPLESDEDEKA FTISSNPYENOKIPTKENVYBAKSYERSVHISL ASVEAQKSHSV APSKPYENMAETIOSPDAN HQXVRAGGNTSEGREYSRSVHISLVASKPPRS HNGNELVESNEKGESSIYISDLNDKRQUJPA ENPCEGGSRARNYEDSVIQSTRAKPQKSVP GEGGGEFKKDGESYYSSBVREYQKARAKKY YEHRISNETSVHISLPFGGTTEPFKGMLYECQ ECGECFAHSSDLTEHQKHEDREKPSGSNNYE WSVIRSLAPTDEQTSYAQEQYAKEQARNICK DFRQFFATSEDLNTNQKIYDQEKSHGEBGOQ KDDPDDKYTSCEDCGLGFVDLTDLTDHOKYV SRKCLVDSREYTHSVHITSISSEYQRDYTGEQ LYECPKCGSSHISSLFEHQRHREQDQLYSW KGCDDGFTALLPRKYBRITAHSEYQRDYTGEQ LYECPKCGSSHISSTLFEHQRHREDDLLSG QMAEALIBGLALTHEPORSOTEELHFECAVCG GSSVPRAEALDHVTVHKNEPYEVGSSYTHST FLIEPFLKGAIPPYSCKDCKSSFHSTVLTKHKE LHLEEBEEDBAAAAAAAAAQEVEANVHYDV VVLRIGGLNVEAAEPEVAAAPPEV EAAPPNGEAGPDGEAAEPEVAAEPEV GDADEPDGVGIEDPEEGEDOEIOVEFFYYDOL HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HECTETFTSSTAFSHLKTHASMIPEPANAFG ECSGYVERASTSTGGANQADEKYFKCDVCGQ HETHIPMDQEVTLUFLIGTTGTTKNLIGCYVCGQ HERPHDPROCECTIVILLIFTGTGTTKNLIGCYVCGQ HERPHDPROCECTIVLLIFTGTGTTKNLIGCYVCGQ HERPHDPROCECTIVLLIFTGTGTTKNLIGCYVCGALDPKH VCNPAIFRLIARQLAKHHAHANGWPKSNL WLKMGKYFSLIPFGADEDINKRFLSDIPSSQI LQEENTWMKELSNLGSVVLCHINDLLCKNNI YNEKGQDVGFDVYSGYVLLVANJGNHRIE FAGVSDDVDSLYPDRELGSQWLRAYLEAV EKGGTVTEREVERLIFTQNOQPALASHFFW GLWALIQAKVSTEEPBLGYAVQFALASHFFW GLWALIQAKVSTEEPBLGYAVQFALASHFFW GLWALIQAKVSTEEPBLGYAVQFALASHFFW GLWALIQAKVSTEEPBLGYAVPKNNINGHNENE FAGVSDDVDSLYPDRELGSQWLRAYLEAV EKGRGTVTEREVERLIFTQNOQPALASHFFW GLWALIQAKVSTEEPBLGYAVQFALASHFFW GLWALIQAKVSTEEPBLGYAVCHINDLLCKTINI YNEKGCOVGFTDVTSCUCTFTTTTTTTTTTTTTTTTTTTTTT							
GYEKSVIHSGFTESOKSHTITRPLESDEDEKA FITSSNPYENGKIPTKENVYESYERSVIHSL ASVEAGKSHSVAGPSKPKVMAESTIGSPANN HQRVRAGGINTSGGRYSKSYERSVIHSL ASVEAGKSHSVAGPSKPKVMAESTIGSPANN HQRVRAGGINTSGGRYSKSYKSYESVIHSL BINGNELVESNEGGESSIYISDLNDKROKIPAR ENPCEGSKNRNYEDSVIQSVFRAKPKK GEGGGFKKSDGFSVFSSNYREYQKARAKKK YIEHRSNETSVIHSLPFGEQTFRYRGMLYEQ GEGGGFKKSDLTEHQKIHDERKPSGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEKSFGGSRNYE WSVIRSLAPTDPQTSV AQEQYAKEQARNKCK DFRQFFATSEDLNTNOKIYDEVIGOLYSK KGCDDGFTALLPMKPRRNRAAERNPALAGSA RCLLCQGGFIRISSLIPETDPUQGSSDMEDPQ KDDPDRIVTECEDGCLGFVEGSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQFVEANVIVPQ VVLRQGLNVEAAEPFVEAAEPFVEAAEPFVACAAEPFV EAAEPNGEAGPDGGAAEPIGGAAEPTGAAEPFV EAAEPNGAAGPDGGAGEDPEERGEAGEPEG GDADEPFDGVGIDPEEGEQQEIQVEEPTYDC HECTETTSTSTAFSEHLKTHASMIPEPANAFG ECSGYIBRASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARRJONTHTG GRPRSSSDNRNFLREERGLSSAAVGTRIGNSA ASRSPAARPVAPVAPVAVVAVAVAVAVAMANNTHV PPOSSEVPKLNVTVQNQDEEHGCGALSLQ HLRPHVDPQGEVTLQLFTDGTTMKLIGGVVGN TMEDVALVRIVQNOKTFLLVDRDEEVXSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIFKSNL WLKMGKYFSLIPTGFADEDINKFILSDIPSSQI LQEEMTWMKELISNLGSFVVLCHNDLLCKNII YNEKQGDVCPIDYFSGYNTLAVBIGNHENE PAGVSLVDVS1, YPDRELQSQWLRAYLEAYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVEKEEVELLIFQVNALASYK EFKGFGTSTVERKEEVELLIFQVNALASYK EFKGFGTSTVERKEEVELLIFQVNALASYR LIMBONALAGRISTBERQHLIGGGLGFE FFFTSAKDINNYKQTFERLVDILICKMSSCLGF FFFTSAKDINNYKQTFERLVDILICKMSSCLGF FFFTSAKDINNYKQTFERLVDILICKMSSCLGF FFFTSAKDINNYKQTFERLVDILICKMSSCLGF FFFTSAKDINNYKQTFERLVDILICKMSSCLGF FFTSAKDINNYKQTFERLVDILICKMSSCLGF FFTSAKDINNYKQTFERLVDILICKMSSCLGF FFTS			<u> </u>				! " " "
FITISSINPYENQKIPTKENVYEAKSYERSVIISIL ASVEAQKSISVAGPSKPK VAMAESTIGSEDAM HQRVRAGGINTSGRE SYSSVIIISIL VASKIPRS HNGNEL VESNEKGESSIYISDLINGKRQKIPAR ENPCEGGSKNRNYEDSVIQSVPRAKPQKSVPA GEGGGEFKKDGEFSVFSSNVREYQKARAKKK YIEHRSNETSVIISLIPGGGTFTPREGMLYEQQ ECGECFAHSSDLTEHQKHIDREKPSGSRNYE WSVIRSLAPTDOTISVAQEQVAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGESGGE NTDGETHSSETHQGTIETBQPUQGSDMEDPQ KDPDDKIYECEDGGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHISTQPQLYSM KGCLDGFSHTHSVITHISTQPQLYSM KGCLDGFSHTHSVITHISTQPQLYSM KGCDGFALIPMERPRENAENPALAGSA IRCLLCQQFIHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEPQRSQTEERLFECAVCG ESSVIPABLADHVTVHKNEPYEYGSSVTHTS FLTEFLKGAIPFYECKDCGKSFHISTVITHKE LHLEEEEDEAAAAAAAAQQVEANVHVPQ VVLRIQGLNVEAAAPEVEAAAPEVE EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAAPEVEAAAPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAAAPEVEAAAFPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPAAAAAAAAAAAAAAAAPEVEAAAFPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAEFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAAFPEVEAAEPEV EAAEPNGEAAGPDGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			1				
ASVEAGKSHISVAGPIKPKVMAESTIGSEPAIN HQRVRAGGINTSGGRYSRSKIVISLI VASKPPRS HNGNELVESNEKGESSIYISDLNDKRQKIPAR ENPCEGGKNRNYEDSVIQSVFRARPKK GEGGSFKKDGFFSVFSNVREYQKARAKKK YIEHRSNETSVIHSLIPFGEQTFRIPRGMLYEQQ GEGGFKKSDLTEHQKHIDBIREKPSGSRNYE WSVIRSLAPTDPQTSVAQEQYAKEQARNICK DFRQFFATSEDLNTNQKIVQDSKHGEESQGE NTDGEETHSEETHIQGETIEDPVIQGSDMEDPQ KDDPDKIYTECEDGCLGFVIDTDHOKVH SRKCLVDSREYTHSVHITHSISEYQRDYTGEQ LYECPKCGSFHISSALFEHQRIHEQDQLYSM KGCDDGFALLPMFRRNARAERNALAGSA RCLLCQGGFHISSALNEHMRLHHEDDLLEQS QMAEAHIPCLALTEP(RSQF)ETERLFECAVCG ESFVNRALADHVTVHKNEPYEYGSSYTHTS FILTEPLKGAJPFYECKDGGKSFHSTVLTKHKE LHLEEBEDEAAAAAAAQEVEANVHVPQ VVLRIQGLNVBAAEPEVEAABPEVAAPPEV FAAEPNGEAGPDGEAAEPIGEAGOPNGEAE QPNGDADEPDGAGIEDPEERAEPGEAAEPE GDADEPDGVGIEDPEBEGDEPGQVEFFYYDC HECTETTTSTAFSEHLKTHASMIFEPANAFG ECSGYBERASTSTGGANQADEKYFKCDVCGQ LYNDHLSLARRIQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSDNTNFRLRERGISSAAVQTRIGNSA ASRRSPAARPPVAPPALPRGRPGTEGSTSLS APALVVAVAVVAVVAVAMAMNTHY PPGSPEVPKLNVTVQDQEBRCCEGALSLLQ HLRPHWDPQEVTLQLFTDGTTNKLIGGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFTQGEALSLLQ HLRPHWDPQEVTLQLFTDGTTNKLIGGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFTQGEALDPH VCNPAFFELLARQLAKHAHAHNGWIFKSNL WLKMGKYTSLIPTGFADEDINKRFLSDPSQI LQEEMTWMKERLSNLGSPVCLEHNDLLCKNII VNEKQGDVGTDYEYSGYNVLADIGINEHNE PAGVSDVDVSLYPPBELQSOWL RAYLEAYK EFKGGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEFTGFLGYAIVRFNQYFKM KPEVTALKVPB FAGVSDVDVSLYPPBELGOWL RAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEFTGH.GYAIVRFNQYFKM KPEVTALKVPB FAGVSDVDVSLYPPBELGOWL RAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEFTGH.GYAIVRFNQYFKM KPEVTALKVPB FAGVSDVDVSLYPPBELGOWL RAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEFTGH.GYAIVRFNQYFKM KPEVTALKVPB FAGVSDVDVSLYPPBELGONDERVSLIDIGSKVGK KERIKLGJUNDTAGGGERYRTITAYYRGAMGFI LMYDITNEESSNAVQDWSTOKTTYSWDDAQ VILVONKCDMEDERVISTERGQHLGFOLGFE FFETSAKDININKYGTFERLGPILGOLGCLAGEF FFETSAKDININKYGTFERLFPPPQPNCAC TORPWNSRVDDFVVNLHELSKGHISPAKDTS							,
HQRVRÁGGNTSGGRYSRSVIHSLVÁSKPRS HNGNEL/ESNEGGSSIYISDLNÖKROKIPAR ENPCEGGSKRAMPEDSVIOSVERAKPOKSVP GEOSGEFKKDGEFSVPSSNVREYQKARAKK YIEHRSNETSVIHSLPFGGTFERRGMLYECQ ECGECFAHSSDL TEHQKHIDERKPGSGRYSP WSVIRSLAPTPOGTSVAGEQVAKEQARNKCK DFRQFFATSEDLNTNOKIYDQEKSHGESGGE NTDOEETHSSETHQGTIEDPVQGSDMEDPQ KDPPDKIYECEDCGLGFVDLTDLTDHOKVH SRKCLVDSREYTHSVIHTISLSYQDTYTGEQ LYECPKCGESFHISSFLFEHQRHEDPUQGSDMEDPQ KDPDDKIYECEDCGLGFVDLTDLTDHOKVH SRKCLVDSREYTHSVIHTISLSYQDTYTGEQ LYECPKCGESFHISSFLFEHQRHEDPUQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHOKVH SRKCLDSREYTHSVIHTISLSYQDTYTGEQ LYECPKCGESFHISSFLFEHQRHEDPUQGSDMEDPQ KODDEDRIYECRDCGREYTHSVIHTISLSTPELQTBLTTHSVIHTISLSTPELQTBLTHOKVH SRKCLDSREYTHSVIHTISLSTPELQTBLTHOKVH SRKCLDSREYTHSVIHTISLSTPELQTBLTHOKVH SRKCLDSREYTHSVIHTISLSTPELQTBLTHOKVH SRKCLDSREYTHSVIHTISLSTPELQTBLTHOKVH KGCDOFFHISALNEHMALHEDDLLEQS QMALEAIHPGLALTEFQRSQTEERLFELQX QMALEAIHPGLALTEFQRSQTEERLFELQX UTSCHLENGER ESSVIPABLADHVTVHKNEPYEYGSSYTHTS FLIEFLKOAIFPYECKDCGSFHISTVLTKIKE LHLEEEEDBAAAAAAAAACQFVEANVHVPQ VVLRIQGLNVBAAEPEVEAABPEVEAABPEV EAAEPNGEAGPDGEAAEPEGAGOPNGEAE QPNGDADEPDGVGIGDFEGGEGOGPNGEAE QPNGDADEPDGVGIGDFEGGEGOGPNGEAE QPNGDADEPDGVGIGDFEGGEGOGPNGEAE QPNGDADEPDGVGIGDFEGGEGOGPNGEAE QPNGDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGAGPNGEAEPE GDADEPDGVGIGDFEGGEGOGPNGEAEPE GDADEPDGVGIGDFEGGAGPNGEAEPE GDADEPGGVGIGDFEGGAGPNGEAEPE GDADEPGGVGIGDFEGGAGPNGEAEPE GDADEPGGVGIGDFEGGAGPNGEAEPE GDADEPGGVGIGDFEGGAGPNGEAE QPRGDADEPGGGAGPDGAAEPE AARDONAAG GDARRGGVGFTGGGASTLIC UTSCHLEGGGAGPAGATAGAG GDARRGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG							l ,
HNGNELVESNEKGESSINISDLNDKRQKIPAR ENPACEGGSKRNYNEDSVSYFRAKPQKSVY GEGSGEFKKDGEFSVPSSNVREYQKARAKKK YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ ECGECFAHSSDLTEHQKIHDREKFSGSRNYE WSVIRSLAPTDPQTSYAGEQVAKEQARNKCK DFROFFATSEDLNTNOKIYDQEKSHGEESQGE NTDGETHSEETHGQETTEDPVIQGSMEDPQ KDPPDKIYNGEEDGGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHSISEYQDAYTGEQ LYECPKCGESFHISSTLFHQRITEGDULYGSW KGCDDGFIALIPAKPRRNRAAERNPALAGSA IRCLLCQQGFIHSSALNEHMRLHREDDLLEQS QMAEEAHIPGLALTEPGRSGTEERLFECAVCG ESFVNPAELADHVTVHKNEPYSYGSYTHTS FLTEPLKGAIPFYECKDCOKSFHISTVLTKHKE LHLEEGEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVGASTEVETLFCAVCG ESFVNPAELADHVTVHKNEPYSYGSYTHTS FLTEPLKGAIPFYECKDCOKSFHISTVLTKHKE LHLEEGEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVGAAEPEVGAAEPEV EAAEPNGEAEGPPDGAGLOPPGGAGOPPGGAE QPPNGBADEPDGAGLOPPEGAGOPPGGAGOPPGAE GPRSSSDINNFILERERAGESSAAVQTRIGINSA ASRRSTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GFRSSSDINNFILERERAGLSSAAVQTRIGINSA ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS APALVVVAVVVVVVVSAVAWMANNTHV PPGSPEVPKLNYTVQDQEEHRCREGALSLLQ HLRHWDPDGVTIQLFTDGTTNKLIGCVVON TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGGAPQLVCTFNNGLCVEFIQGEALDPICH VCNPAIFFILIARQLAKHAHIAHNOWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEMTWMKELISNLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYNIVCHNDLLCKNII YNEKQGDVQFIDVEYSGYNYLLAYDIGNIFINE FAGYSDVDYSLYPDRELQSWURAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYNIVFKINYFKN EKRIKLQTWDTAGGERYKSTUTTAYYRGAMGFI LMYDITNEESPNAVQDWSTQIKTYSWDNAQ VILVGNKCOMBEBRVISTEGOHLGFGOLGFE FFETSAKDNINVKQTFERLVDIICDKMSSSLET DPAITAAKQNTBLLKETPPPQPNCAC 1344 2694 A 10395 2 4136 DPRPWNSRVDDFYVTILISKGKHISPAKDITS						!	l
ENPECEGGSKNRNYEDSVIQSVFRAKPQKSVPQ GEGSGEFKKNGEFSVPSKPYQKCRAKKK YIEHRSNETSVIHSLIPFGEQTFRPRGMLYECQ ECGECFAHSSDLITEHQKHIDTEKFSGRNYE WSVIRSLAFTDPQTISYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQCKSKGESQGE NTIGGETHISEETHQGETIEDPVIQGSDMEDPQ KDPPDKIYFECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSWHTHSISEYQRDYTGEQ LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKFRRNRAFRALAGSA IRCLLCQGGFIHSSALNEHMRHREDDLLEQS QMAEEAIIPGLALTFFQRSQTEERLFECAVCG ESFVNPAELADHVTVIKNEPYEYGSSYTTHS FLITEPLKGAJFFYSCKDCGKSFIHSTVLTKHKE LHLEEEEEDBAAAAAAAQEVEANHVPQ VVLRQGLNYEAAEPSVEAAEPEVGAAEPSV EAAEPNGEAEGPDGEAAEPIGGAQPMGEAE QPNGDADEPDCVGIEDPEGEEDQEIQVECYPYDC HECTETTISTSTAFSHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHISLARHQNTHTOI 1342 2692 A 10393 2 1350 GRPFSSSDNRNFLERAGLSSAAVQTEIGNSA ASRRSPARPPVPAPPALPRGRPGESTSLS APAVLVVAVAVVVVVSAVAWAMANTIHV PPGSFEVFKLINTVQDQEEHRCREGALSLLQ HLRPHWDPGSVTLQLTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEKSFRVVL QAHGCAPQLVCTFNNGLCVSFIQGEALDPKH VCNPAIFRILARQLAKHHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDDINKRFLSDPSSQI LQEMTWMKELSNLGSFVVLCHNDLLCKNII YNEKQGDVQFIDVEYSGYNYLAYDIGNHENE FAGVSDVDYSLYPDRELGQGWLADPKH VCNPAIFRILARQLAKHHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDDINKRFLIGNSSVGK TSFLFRYADDSSTSAFVSTVGIDFKVTVFKN EKRIKLQIVDATAGGRETYSLGYWLATALAYK EFKGFGTEVTEKEVEILFQVNGFALASHFPW GLWALIQAKKYSTEEPFLGGWLATAHAYK EFKGFGTEVTEKEVEILFQVNGFALASHFPW GLWALIQAKKYSTEEPFLGGWLATAHAYK EFKGFGTEVTEKEVEILFQVNGFALASHFPW GLWALIQAKKYSTEEPFLGGWLATAHAYK EFKGFGTEVTEKEVEILFQVNGFALASHFPW GLWALIQAKKYSTEEPFLGGWLATAHAYK EFKGFGTEVTEKEVEILFQVNGFALASHFPW GLWALIQAKKYSTEEPFLGGWLATATYYRGAMGFI LMYDITNEESPHAVQDWSTQIKTYSWDNAQ VILVGNKCOMDERKVISTRUTTAYYRGAMGFI LMYDITNEESPHAVQDWSTQIKTYSWDNAQ VILVGNKCOMDERKVISTRUTTAYYRGAMGFI LMYDITNEESPHAVQDWSTQIKTYSWDNAQ VILVGNKCOMDERKVISTEQHLJESKGHIFFAKKDITS DPAITAAKQNITALKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DPRPWNSRVDDFYVTLILLISKGHISPAKKDITS							
GEGSGEFKNDGEFSVPSSNVERYOKARAKKK YTEHRSNETSVHSLPFGETPRRGMI, YECO ECGECFAHSSDLTEHQKIHDREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFROFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGETHSEETHGQETTEDPVIQGSDMEDPQ KDPPDKIYDCEDGCIGGFVDL.TDL.TDHQKVH SRKCLVDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNAAFRNPALAGSA RGCLLCGQGFHRSSALNEHMRLHREDDLLEQS QMAEEAIJPGLALTEFGRSFHSTVLTKHKE LHLEEBEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEFEVE EAAEPNGEAAGPPVEKADEFSHISTVLTKHKE LHLEEBEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEFEVE GDABEPDOVGEDPFEGEGFGAEFPE GDADEPDOVGEDPFEGEGVEFPYDC HECTETFTSSTAFSEHLKTHASMIFPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTO PCSSYJERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTO PCSSYJERASTSTGGANQAMDKYFKCDVCGQ ARRISPARARPPVAPAPLERGGTGESTSLS APAVLVVAVAVVVVVSAVAWAMANYHH PPGSPEVPKLNNTVQDQEDEHRCREGALSLQ HLRPHWDPQEVTLQLFTDGTTNKLIGCYVGN TMEDVVLVRIYGOKKTELLVDRDEEFKSFRVL QAHGCAPQLVCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAIHAHNGWIFKSNL WIKMGKYFSLIPTGFADGEVASFPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFTDYEYSGYNVLAYDIGNHFNE FAGYSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFTQVNQFALASHFPW GLWALIQLAKYSTEEFDELGVAIVFRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIIGNSSVGK TSSLFRYADDSSTSAFVSTVGIDFKKVTVFKN KPEVTALKVPE EKRIKLQIPUDTAGGGRYRYTTITTAYYRGAMGFI LMYDITNEESSPNAVQDWSTQIKTYSYDDNAQ VILVGNKCDMDERVISTERTTYYRDDNAQ VILVGNKCDMDERVISTERTYSYDDNAQ VILVGNKCDMDERVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDMDEBRVISTERTYSYDDNAQ VILVGNKCDM				· ·			l
BEGGECFAHSSDLTEHQKHEDREKPSGSRNYE WSVRSLAPTDPQTSAQDYAKEQARNACK DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGEETHSEETHGQETIEDPYIQGSDMEDPQ KDPPDDKIYECEDGCIGFVDLTDLTDHOKYH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEG LYECPKCGSSHHSST,FEHQRHECDQLYSM KGCDDGFTALLPMRPRRAAERNPALAGSA RCLLCQQGFHSSALNEHM,REDDLLEQS QMAEEAIIPGLALTEPQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPPYGSSYTHTS FLIFPLKGAIPFYECKDCGKSFHSIVLTKHKE LHLEEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGLNYBAAEFEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGCAACPEV EAAEPNGEAEGPDGEAAEPIGCAACPEV EAAEPNGEAEGPDGGAGIEDPEGAEEPCKAEPEV GDADEPDGVGIEDPEGAEPFYDC HECTETTTSSTAFSEHLKTHASMIIFPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHONTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPAPLFRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIEN PPGSPEVPKLNVTVQDEFIRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNGLCYEFIQGEALDFKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYTSLIPTGFADEDINKRLSDIPSQO LQEEMTWMREILSNLGSPVVLCHNDLLCKNII YPEKQGOVQFIDYFSGVNTLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTLEPFJCAYAVRNGYFKM KPEVTALKVPE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW CLWALIQAKYSTLEPFJCGYAVRRNGYFKM KPEVTALKVPE FAGVSDVDFSTAAVSTVGIDFKVTTVFKN KPEVTALKVPE EKKRIKLQINDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQKYTSWDNAQ VILVGNKCOMBEDERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQKYTSWDNAQ VILVGNKCOMBEDERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTGKYTTSWDNAQ VILVGNKCOMBEDERYSTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCCKMBESLET DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQPNCAC DPAITAAKQNITILKETPPPPQENCAC DPAITAAKQNITILKETPPPPQENCAC DPAITAAKQNITILKETPPPPQENCAC			j .]]	
WSVIRSLAPIDPQISYAQEQYAKEQARNIKCK DFRQFFATSEDLATINQUYQEKSHGEESQGE NTDGETHISETHGQETTEDPVIQGSSMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHGKVH SRKCLVSREVTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHECDQLYSM KGCDDGFTALLPMKPRRNRAAERNPALAGSA IRCLLCQQGFHISSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEQRSQTEERLFECAVCG ESFNYPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHSTVLTKHKE LHLEEEEDFAAAAAAAQFVEANNYHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAQQPNGEAA QPNGGADEPDGAGIEDFEAEEPEGKAEPE GDADEPDGVGIEDPEEGEDOBIQVEEPYYDC HECTETFTSTSTAFSEHLTASAMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARRQNTHTG GRPFSSSDNRFFLRERRGLSSAAVQTRIGNSA ASRRSPAARPYPAPPAPLFRGRPGTEGSTSLS APAVLVAVAVVVVVVAVAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYYON TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAEFLIARQLAHAHAHNGWIPKSNIL WLKMGKYFSLPTGFADEDINKRPLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKGGOVQFIDYEYGGNVLAYDLGHYFNG GLWALIQAKYSTLEFDFLGYAIVFNQYFKM KPEVTALKVPB 1343 2693 A 10394 102 839 PEAQTSAVLAREKGGHLPTIMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN KPEVTALKVPB 1341 2694 A 10395 2 4136 DRIPWNSRVDDFVTNLETSRGQHLGEGOLGFE FFETSAKDNINVKQTFERLVDLICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC			l '				YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ
DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGETHSEETHGQTIEDPVIQSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGSSTHSSTEPHQRIHEQDQLYSM KGCDDGFTALLPMKPRRNRAARRIPALAGSA IRCLCGGFIHSSALNEHMRIHREDDLLQS QMAEAIIPGLALTEFQRSQTEERLFCCAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLITEPLKGAIPFYECKGSFHISTVLTKHKE LHLEEEEDERAAAAAAAAAQEVEANVHVPQ VVLRIQGI,NVEAAEPEVAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEGEDDGEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARIQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLREAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDEFICKGEALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCVVGN TMEDVVLVRVIGNKTLURDEDEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKGGOVOFIDVEYSGVNVLAYDIGNIFINE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTLEPTGADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKGGOVOFIDVEYSGVNVLAYDIGNIFINE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTLEPTGFLGYAIVRRNQVYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSFLIFRYADDSFTSAVVSTORYTKMYNYFKM KPEVTALKVPE EKRIKLQINDTAGGREYSTGTTATYYRGAMGFI LMYDITNEESFNAVQDWSTGKYTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTGKYTTSWDNAQ VILVGNKCOMBEDERVSTSTERGQHLGEGLGFE FFETSAKDNINVKQTFERLVDLICCKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSKVDDFVTNLHLESSKGHLSPAKDTS		1					-
MTDĞEETHSEETHGQĒTIEDPVIQGSDMEDPQ KDDPDDKIYECEDGGIGFVDLTDHOKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESPHSSFLFEIQRHEQDQLYSM KGCDDGFIALIPMKRNRAAERPALAGSA IRCLLCGQGFHSSALNEHMRLHREDDLLEQS QMAEAIIPGLALTEFQRSQTEERLFECAVCG ESFYNPAELADHVTVHKNEPYEYGSSYTHTS FILTEPLKGAIPFYECKDCGKSFHSTVLTKHKE LHLEEEBEDEAAAAAAAAQEVEAAHVHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPEVGAAEPEV GDADEPDGVGIEDPEGGEOPQRVEEPYYDC HECTETFTSSTAFSEHLKTHASMIFPEANAFG ECSGYBERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG ECSGYBERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG BECSGYBERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIV PPGSPEVPKLNVTVQDQEBHRCREGALSLQ HLRPHWDPQEVTLQLFTDGTTNKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEPIQGEALDPKH VCNPAHFBLIARQLAKHAHHAHNGWPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELSNLGSPVVLCHNDLLCKNII YNEKQGDVÇFTDYSGYNYLAYDIGNHFINE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFTEVTEKEVEELFIQVNOPFALSHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTJCKTYSWDNAQ VILVGNKCDMEDBERVISTERGQHLGFOLGFE FFETSAKDNINVKQIFFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC	1						
KDDPDDKIYECEDCĞLGFVDLTDLTDHOKVE SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECFKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRAERNPALAGSA IRCLLCGQGFIHSSANEHMRLHREDDLLCGS, QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVIHKDEPYEYGSSYTHTS FLTEPLKGAIPPYEVESVTHTS FLTEPLKGAIPPYEVESVTHTS FLTEPLKGAIPPYEVENCOGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAAPEVEAACPEVE EAAEPPUGAAPEVEAACPEVE EAAEPPUGAAPEVEAACPEVE EAAEPPUGAAPEVEAACPEVE EAAEPPOGAGEDPEERGEPGGAEEPE GDADEPPDGVGIEDPEERGEDQEIQVEEPYYDC HECTETTSTSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCOQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNNIPLRERAGISSAAVQTRIGNSA ASRRSSNNIPLRERAGISSAAVQTRIGNSA ASRRSSAARPAPVAPALPRGRPGTEGSTSLS. APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHVDPQEVTLQLFTDGTINKLIGCYYON TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDSLYPDRELGSQMLRAYLEAYK EFKGFGTEVTEKEVEELFIQVNQFALASHFFW GLWALQAKYSTEEPDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFRAVQDWSTJKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPPQNCAC		· ']	į			
SRRCL/DSREYTHSVIHTHSISEYQRDYTIGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNAAERNPALAGSA IRCLLCGQGFHSSALNEHMRLHREDDLLEQS QMAEAHIPGIALTPGRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FILTEPLKGAIPFYECKDCGKSFHSTVLTKHKE LHLEEEEDEAAAAAAAAAQEVEAAHVHVPQ VVLRIQGLWEAAAPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAEPIGEAGQPNGGAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGEDOEIQVEEPYYDC HECTETFTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCOQ LFNDHLISLARHQNTHTG GRPRSSSDNNRNLEREAGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVAVAWAMANNYIHV PPGSEVPKLNVTVQDQEEHRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDFKH VCNPAIFFELIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDVEYSGYNVLAYDIGNHENE FAGVSDVDVSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSTSAFVSTVGIDFKVNTVFKN EKRIKLQIWDTAGGERYRITITAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDBERVISTERGGHLGEQLGFE FFETSAKDNINVKQTFERLYDILCDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVNLHLISSKGHISPAKDTS							
LYECPKCGESFIHSSFLFEHQRIHEODQLYSM KGCDDGFIALLPMKPRNRAAERNPALAGSA IRCLLCQQGFIBSALNEHMRLHREDDLLEQS QMAEEAIIFGLALTEPGRSQTEERLFECAVCG ESFVPNAELADHVTVHKNEPYEYGSSYTHST FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAEPIGEAGOPNGEAE QPNGDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETITSSTAFSEHLKHASMIFEPANAFG ECSGYIERASTSAFSHLKHASMIFEPANAFG ECSGYIERASTSAFSHAKHASMIFEPANAFG ECSGYIERASTSAFSHASHAWAANAANYHIV PPGSPEVPKLNVTVQDQEEHRCREGALSLAQ HLRPHWDPQEVTLQLFTDGITNKLIGCVVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIGGEALDPKH VCNPAIFRLIARQLAKHAIHAHNNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDPSSQI LQEEMTWMKEHLSNLGSPVVLCHNDLLCKNII YNEKQGDVGFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFLQVNOFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVATREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLISSKGHISPAKDTS							7
KGCDOGFIALL PMKPRRNŘÁAERNPÄLAGSA IRCLLCQGGFIASLA PMKPRRNŘÁAERNPÄLAGSA IRCLLCQGGFIASLA NEHMRLHREDDLLEQS QMAEEAIIBGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEBEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAGPEVEAAGPEVE EAAEPNGEAEGDGEAGEPGEAGOPNGEAE QPNGDADEPDGGAEAFPIGEAGOPNGEAE QPNGDADEPDGGAEAFPIGEAGOPNGEAE QPNGDADEPDGGIEDPEEGEGQEIQVEEPYYDC HECTETTSTSTAFSEHILKTHASMIFEPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG GFRFSSSDNRFILRERAGISSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS AAPAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGTINKLIGCYYGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFANGLCVEFIQGBALDPKH VCNPAIFRLIARQLAKIHAHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEVSGYNYLAYDIGANHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FAGYSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE KERIKLQIVDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERQQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICCKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							, , ,
IRCLLCGQGFIHSSALNEHMRLHREDDILLEQS QMAEEAIPGLALTEPQRSQTEERLFECAVCG ESFVNPAELADHVITYHKNEPYEYGSSYTHTS FI.TEPLK.GAIPPYECKDCGKSFIHSTVLTKHKE LHLEEBEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVBAAEPEVEAAEPEVEAAEPEV EAAEPNGBAEGPDGBCAEPIGEAGOPNGBAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETTSSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGISSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFITDGTNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKFLSDIPSSQI LQEEMTWMKEILSNLGSPVLCHNDLLCKNII YNEKQDVQFIDYEYSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQWQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLIPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN KPEVTALKVPE 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					•		
QMAERAJPGLALTEPQRSQTEERLFECAVCG ESFVPAELADHVTVHKNEPYEYGSSYTHTS FLIEPLKGAJPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAGPDGEAEPGGAEPGGAEPE GDADEPDGAGIEDPEGEGDQEJQVEEPTYDC HECTEITTSTAFSEHLKTHASMIFFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMNYHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLLCHTDGITNKLIGCYVGN TMEDVVLVRYIGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYPSLPTGFGADEDINKRFLSDPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKGDVVGIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETFDFLGSQWLGSGWLGSGWLGSGWLGSGWLGSGWLGSGWLGSGW							The state of the s
ESFVNPAELADHVTVHKNEPVEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAEPEV EAAEPNGEAEGPDGEAGEPVEAAEPEVEAEPEG GADEPDGVGIEDPEEGEGAGEPEGAGEPPGGDAEPIGAGGPNGEAE QPNGDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPVPAPPAPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTI,QLFTDGTINKLIGGYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFIQVNQFALASHIFFW GLWALIQAKYSTEEDFLGYAIVFRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTOJKIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTIKLETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHILSSKGHISPAKDTS							
FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDDAAAAAAAAAQCVEANVHVPQ VVLRIQGLWAAAFPEVEAAEPEVAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPPO VGIEDPEGEDOEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVVVVVVVVVAVAWMANYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRPLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQJKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPOPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
LHLEEEEDEAAAAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAAEPEVEAAEPEV EAAEPRIGEAGOPNGEAE QPNGDADEPDGAGIEDPEEGEDQEQVEEPYYDC HECTETITISTATSEFILKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSONRFLRERAGI.SSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSFEVPKLNVTVQDQEBHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFIRLIARQLAKIHAIHAHAINGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEVSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAGTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGQKDSSDQNFDYMFKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQGRYRTITTAYYRGAMGFI LMYDITNESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPOPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					ļ		I
VVLRIQGLNVEAAEPEVAAEPEVAEAEPEV EAAEPPUGEAGQPNGEAE OPNGDAEPDGAGEOPDGEAAEPIGEAGQPNGEAE OPNGDAEPDGAGIEDPEERAEPEGKAEPE GDADEPDGAGIEDPEERAEPEGKAEPE GDADEPDGVGIEDPEERAEPEGKAEPE GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETITSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKKFGTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNESSFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVAQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							1
EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEEGAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETTTSSTAFSEHLKTHASMIIFEPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDRRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAJFRLIARQLAKHAIHAHNGWPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNI.GSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLA YDIGNHFNE FAGVSDVDYSL YPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	}						,
QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPTYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRFGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNTVVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCVGVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHENE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC							•
GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVPPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNYTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC							
ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHITG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAIHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGPRGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC							
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNIFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HILRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1342	2692	Α	10393	2	1350	,
PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				'			
TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1						
QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVYLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNIII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					-	•	
LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							•
YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							, ,
GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							· · · · · · · · · · · · · · · · · · ·
KPEVTALKVPE						ľ	
1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					İ	ŀ	
QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1343	2693	Α	10394	102	839	
TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS						' I	
LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					į	ļ	
FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS]					ŀ	
DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							, ,
1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS						}	,
		260					
LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR	1344	2694	Α	10395	2	4136	·
	L						LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR

NO. of NO. of NO. of nucleo n	SEQ I	D SEQID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleoride cotide sequence uence under the corresponding of the corresponding of peptide residue of peptide residue of peptide sequence peptide sequence seq		, ,		, ,			1
Dealing Deal)			1 *
sequence petide greided greide		,					
genice 914 ng to first amino acid residue of peptide residue of peptide peptide of peptide sequence y=1000000000000000000000000000000000000	1			Į.			1
amino acid residue of peptide sequence peptide sequence peptide sequence of peptide sequence	1 -	"""		1		4	, , , , ,
residue of peptide sequence peptide sequence peptide sequence sequ	l acrice	- 1	1	'''			
Poptide sequence		ł					
Sequence nucleotide insertion	1	- 1	l			sequence	
RRLQGKLPELQGVETELCYNVNWTAEALP EFTKKLMWLFGCFULDDWARSSWLFOSS DLLLEVOPRUNSTPTSTNIVSVCRATOLOP DRVETTERYBLSAPHPSAEVEALALATHE MTEQHEPHPIQSTSPSMPEPLNGPINLIGGG LALEKANQELGLALDSWDLDPYTKRFQELC WSTVZAFDLQSNSBHSRHWFKQGLHVT QKLYHSLFSMSTQESSNFNNVLKFCDNSS QKGEWRFLRREDPTRSRFQQQGCGLRHVY AETHNPTIGVCPTSGATTGGRIRDVQCTT RGAHVVAGTAGYCFGNLHIGVTNLPWEDL QYPGNFARRLEVAIEASNGASDYGNKFGEPP LAGFARSLGLQLPDGQARSWNFNGFGGF AGASSQVQQDNTSDLDFGAVGROPEME KMNRVIRACVEAPKGNPICSLHDQGAGGNN NVLKELSDPAGAJITYSEFQLGDPTLANLEIV GAEYQESNALLLRSPNRDFLTHVSARRCPP, CPVGTITGDRINLVDDRECPVRRNQGGDA PTPPTPYDLELEWVLGKNPRKEFFLQNRPP MQPLALPPGLSVBQALEFVLAPVASKR LTNNVDRSVGGLVAQQQCVGPLQTPLADV VVALSHEELIGAATALGEGPVKSLLDPKVA RLAVEALTNLVFALVTDLRDWCCSGNWM WAKKLPGGGALADACEAMVAMALGV VDGGKDSLSMAARVGTETVRAPGSLVISAV VQPDTATYTPDLKHEPGRGHLLVALSPC AFAGNGCQVDVPRVDVISLVLFALEFGL AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE AFAGNGCQVDVPRVDVISLVLFALEFGL LEVQEPDLAQVLKRYRDAGLHCLLEGITTGE BMADAFHLAGFEVWDVTNQDLCSGAGGLD FROVAFVGGGSTSANAVVLEEVGELEALWE TSRQLDRLQAEPRCVAEEERGLBERNGDSY LPTTPFASNYREPGGFSPRVALIREGSNGC EMADAFHLAGFEVWDVTNQDLCSGAGGLD WVGGSPPIDAAEMOPDSQPARPOPDLLTRIN SGRYESRWASVRVQFGPALMLRGMGCAVL VWSAHGGYVAFSSPELQAQIEAGAAATTH RAGAELRRFRRPDTTSLGVCNGCQLLALL WYGGGPPIDAAEMOPDSQPARPOPDLLTRIN SGRYESRWASVRVQFGPALMLRGMGCAVL VWSAHGGYVAFSSPELQAQIEAGAAATTH RAGAELRRFRRPDTTSLGVCNGCQLLALL WYGGGPYBDAAEMOPDSQPARPOPDLLTRIN SGRYESRWASVRVQFGSPRVALIREGSNGC HLAWMPHPERAVRPQWAWMPPPDTLTT PWLQLFNRRNWTLEGSC GRGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSSRTTHFELKHLSSGDLLRDMMLKG GKGTVSRTTHFELKHLSSGDLLRDMMLKG GKGTVSRTTHFELKHLSSGDLLRD		ļ	ļ	l	,		1 •
BETKKLMWLFÖCPLLIDDVARESWLLPGS DILLEVOPRINSTPTSTNIVSVCATAGLOP DRVETTERYRLSFAHPPSAEVAJALATILH MTEQHPHPIGSTPSTNIVSVCATGLOP DRVETTERYRLSFAHPPSAEVAJALATILH MTEQHPHPIGSTPSSMPPENLOPINILGEG LALEKAPQELGLALDSWDLDFYTTKRQELC NPSTVEAFDAGSNSEHSRHWFKQCHIVE QKLVIHSI,FESIMSTQESSNINNVLKFCDNSS IQKEVFELRREDPTRYSRFQQQGCIRHVYV AETHINPTTOVEPSGATTGTTGGRIRDVQCTT RGAHVVAGTAGYCFONLHIRGYNLPWEDL: QYPGNFARFLEVARSNGASDYORKFGEP LAGFARSLGLQLPDGQRERWIKPIMFSGGIG MEADHISKEAFPEDMEVVKYGGPYTRIGVGGAGSON MEADHISKEAFPEDMEVVKYGGPYTRIGVGGAGSON MENTINCHELSDPAGAHYTSSFQLGDFTLNALERI GAEVOGSNALLLRSFNODELTHVARRERP, CEVGTTITGDRRIVLYDDRECPVRNOGGGAGN NVLKELSDPAGAHYTSSFQLGDFTLNALERI GAEVOGSNALLLRSFNODELTHVARRERP, CEVGTTITGDRRIVLYDDRECPVRNOGGGAGN PTPPTPYDLELEWUGKMPREFFLQREP GAEVOGSNALLRATAGEGPVKSLLDPKVA RLAVBALTINLYFALVTDLRDVKCSGNWM WAAKLPGEGALADAGAAVAMALGV VOALSHEELIGAATAGEGPVKSLLDPKVA RLAVBALTINLYFALVTDLRDVKCSGNWM WAAKLPGEGALADAGAAWAAVAALGV VOBGKDSLSMARRVGTETVRAPGSLVISAY VCPDITATVTPLKHFEGRGHLLYVALSFGG HRLGGTALAQCFSQLGEHPDLDLPENLYR FSTGOLLKDRLLCSGHDVSDGGLVTCLE AFAGNCGLQVDVPVRVDVLSVLTAEEPGG LEVQEPDLAQVLKRYRDGLHCLEGHTGG AGPHAMYRVSVNGAVLEEPVGELRALWF STGOLDRLQAEFFGCVAEFEGRLERRMGPSY LPPTPKASVPRPGGPSPRVALLREGGSNGF LPPTPKASVPRPGGPSPAVALLREGGSNGF LPPTPKASVPRPGGPSPAVALLREGGSNGF LPPTPKASVPRPGGPSPAVALLREGGSNGF LPPTPKASVPRPGGPSPAVALLREGGSNGF LPPTPKASVPRPGGPSPAVALLREGGAGLAPLH ADDDONNTEGYPLHPNCSFGGVAGICSCDG BANADAFHLAGFEVWDVTMODLCSGAIGLD RAVMHYQMSVTLKYPIKKLLYHLJVIHLJUKLLL AKMSYGHRLLSHDQVAMPQVWEPYLLS LPSLGLLSFPRNNISVLVLSMISMGIFSILPH YGGMWAGTMASKAURAAVTHF ADDDONNTEGYPLHPNCSFGGVAGICSCDG GROWGFWAGTMASKAGARAGTRAGTDGSDFQHR GVARMPHPERAVRYRWWAWRPPPDTLTT PWLQLFNARNWTLEGGS GGTVSSRITTHFELKHLSSGDLLRDNMLRG GKGTVSSRITTHFELKHLSSGDLLRDNMLRG GKGTVSSRITTHFELKHLSSGDLLRDNMLRG GKGTVSSRITTHFELKHLSSGDLLRDNMLRG GKGTVSRRITTHFELKHLSSGDLLRDNMLRG GKGTVSRRITTHFELKHLSSGDLLRDNMLRG GKGTVSRRITTHFELKH	<u> </u>		 	<u> </u>	Sequence		
DLLLEVGPRINSTPTSTNIVSVCRATGLATALIE DRVETTRRYRIS.FAHPPSAEVALATALIE MTEQHPPIPIOSISSPESMPEPLNGPINILGE LALEKANGEL GLAL DSWIDLDFYTKRFOGEL LALEKANGEL GLAL DSWIDLDFYTKRFOGEL NPSTVEAPDLAGSNSEHSRHWFFKGOLHVU QKLVHSLFSIMSTQESSNPNNVLKFCDNSS IQGKEVRFLRPEDPTRRSRPQQQGLRHVY AETHNPFIGVCPFSCATTGTGGRIRDVGT RGAHVVAGTAG VCFGNLHPGYNLFWEELS QYPGNFARPLEVAIEASNGASDVGNKFGGEN MEADHISKEAPEPGMEVVKVGGPVYTIGS MEADHISKEAPEPGMEVVKVGGPVYTIGS GARSVQVGGNTSDLPGOAVGROPPEME KMMRVIRACVEAPKGPIPCSIGNE GASYQCSONTSDLPGOAVGROPPEME KMMRVIRACVEAPKGPIPCSIGNE GASYGESNALLLRSPNROFLTHVSARERF GAEYGESNALLLRSPNROFLTHVSARERF GAEYGESNALLLRSPNROFLTHVSARERF GAEYGESNALLLRSPNROFLTHVSARERF PROFTPTFVDLELEWVLGKMPRKEFFLQRKF MUPPLALPPGLSVHOALEREVKLIPAVASSR LTNKVDRSVGGLVAQQCVGPLQTFLAGLEN VALSHEELIGAATALGEQPVFRNOGGDA PTPPTTYDLLEBWVLGKMPRKEFFLQRKF MVAKLPGEGAALADACEAMAVAMAALGV VDGGKDSLSMAARVGTETVRAPGSLVILY VVALSHEELIGAATALGEQPVSLLDPEVA RLAVAEALTNLVFALVTDLRDVKCSGMVAV VDGGKDSLSMAARVGTETVRAPGSLVILY VDGGKDSLSMAARVGTETVRAPGSLVILY RAGGELATALGCFSQLGEHPPDLDLPENLYR FSITGGLLKDRLLSGBIDVSDGATVLLSEAGEN FSITGGLLKDRLLSGBIDVSDGATVLSVLAFAEPEGL LEVCEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVLEPVGELRALWE TSFQLDRLQAEPRCVAEEBRGLRERMGFSYSY LPPTTPRKASVPREGGSSRVALUREGGSNOE EMADAFHLAGFEVWDVTMQDLCSGAIGUTULL RAGGELARGFRKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARGLLLRIN SGRYESSWASVRVOPGPALMLRGGGSNOE EMADAFHLAGFEVWDVTMQDLCSGAIGUTULL RAGGELARGFRKARPBYPTLLTH RAGAELRRFKKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARGLLLRIN SGRYESSWASVRVOPGPALMLRGMGSOLV VWSSHGEGVYAFSSPELQAGEARGLAPLH ADDDGNPTEQYPLNPNGSFGGVAGIGSCDL ADDDGNPTEQYPLNPNGSFGGVAGIGSCDL LAVMPHERALRSDLLRINGNBGAPL PSILGLLSFFRNNISYLLSMSMGLFSBL LPSILGLSFFRNNISYLLSMSMGLFSBL TJ845 2695 A 10396 65 642 GVGFGWAGTMSARAGFRAGTDGSDFQIR RVAMHVQMSVTLKYLRILSSGDLLRDMLBGEN TYGGDLLIGEPUQREDOKPETVIKLLKANYPL VGGSMEMFBAAQQLVRHCKAXRELFGFSAV MYLVLUAQVYAWQLYYSKKLLDSWFTS QEKKHK TYGGDLLTGEPLQREDOKPETVIKLKLRSWPY VGGDDLLAGFILOGGLURPDLWARLDRAFQUTDTVIKL NVPPEVIKGRLTARWHPASGRGR TYGGDDLAGETURLBURYPPF TVGIDDLTGEPLQREDOKPETVIKLK	1]	j	Ì	j		
DRVETTRYRISTAHPPSAEVEAIALATUR MTEGHPPHPIGSFSERSMENDGINILGEG LALBIKANOEL CLALDSWOLDFYTKEFOELL NPSTVEAFDLAQSNSEHSRHWFFKQCHVV QKLVIISLIFISIMSTQESSNFNNVLKFCDNSS IQGKEVERLRPEDFTRPSRQQQQCLRIVY AETHNEFTGVCPFSGATTGTGGRIRDVQCTO RGAHVVAGTAGVCFGNLHIPGYNLPWED JQYPGNFARPLEVAIBASINGASDYGNKFGEP LAGFARSLGLQLPDGQREWIKPIMFSGGIG MEADHISKEAPEFGMEVVXCPY GAASSVQVGGDNTSDLDFGAVQRGDFBW KMMRVIRACVEAPKGMPICSLHDQGAAGM NVLKELSDPAGAHIYTSRFQLGDPTLNALBY GAEYQESNALLLRSPNRDFLTWASBECP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPPTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPPTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPPTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPPTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPPTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTPTVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTPTVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTPTVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTTPVDLLEWVLGKMPKABERCP, CEVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTTPVDLLEWFGGGHHPDLDLPENLVRR WAAKLPGEGAALADACEAMVAWAAALGV VOGGKDSLSMAARVGTETVRAGSLSVILEY VOGGKDSLSMAARVGTETVRAGSLSVILEY FSTGCLLKRDLLCSGHIVSDGCVTCLLEN AFAGNCGLQVDVPVPRDVLSVLEAEEER, FSTTGCLLKDRLLCSGHIVSDGCVTCLLEN AFAGNCGLQVDVPVPRDVLSVLEAEER, FSTTGCLLKDRLLCSGHIVSDGCVTCLLEN AFAGNCGLQVDVPVPRDVLSVLEAEER, FSTTGCLLKDRLLCSGHIVSDGLATCLLELGGTTG AGPHAMVRVSVNGAVVLEEPVGELRALWF TSFQLDRLQAERFCVAEEERGMPSYY LPPTPFKASVPREFOGPSRVAULREGGSNGC EMADAFHLAGFRVDVTMGSGGCVAGLCSCGL LEVVGFPDLAFVLSSGFLGRAGHLAUL AGMSVGHLRLLSGGLVCSAGGL EMAGAELRFFRKRPDTTSLGVGCCLLALL VVGGDPPEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGDPSRVAULREGGSNGC EMADAFHLAGFRAVBOWAWARPPFDTLTT PWLQLFINARNVTLEGSSC GVGFWAGTMARSAFAGFRAAGTDGSDFQHR RVAMHYGMSVTLKVLSHMSMGLFSHD VGSMEMFFAAQOLVRHGKAYRFFFGFSAV MYLVLVLAVQVHAWQLVYSKKLLDSWFT QEKKIK DDFFYEGFQSAAMGASARLLRAVINGARGG GKGTVSSRTTHFELKHLSSGDLLRDMRGG GKGTVSSRTTHFELKHLSSGDLLRDMRGG GKGVLKKW NVPFEVLKGLTARWHPASGRVYNIEFNPF TVGIDDLTGEPLIQREDDKPETVIKLLKAYN NVPFEVLKGLTARWHPASGRC GKGVLAKAFILOGGKLWPVV NSPEVLKGLTARRWHPASGRVYNIEFNPF TVGIDDLTGEPLIQRED		1		ļ.	l		
MTEQHPPIPIOSISPESMPEPLNOPINILGE LALEKANQELGLAL DSWIDLPYTKRFQELG LALEKANQELGLAL DSWIDLPYTKRFQELG NPSTVEAPDLAQSNSEHSRHWFFKGQLHVY QKLVHSLFSIMSTQESSNPNNVLKFCDNSS IQGKEVERLRPEDPTRYSRFQQQQGLRHVY AETHNFPTGVCFPSGATTGTGGRIRDVQCT RGAHVVAGTAGVCFGNLHIPGYNLPWEDLS QYPGNFARPLEVAIEASNGASDYGNKFGGEP LAGFARSLGLQLPDGQRREWIKPMFSGGIG MEADHISKEAPEFGMEVVKVGGPVYNIGS GAASSVQVGGDNTSDLDFGAVQRGDPEME KMRVIRACVEAPKGNPICSLHDQGAGGNO NVLKELSDPAGAUTYSRFQLGAGATYSRFQLGAGATY GAEYQESNALLLSEPNRDFLTHVSAREGCP, CFVGTTITGDRRIVLVDDRECPVRRNGQGDA PTPPPTPVDLELEWVLGKMFRKEFFLQRKFF MLQFLALFPGLSVHQALERVLRLPAVASKE LTNKVDRSVGGLVAQQCVCPLQTPLADL VVALSHEELIGAATALGEQFVKSLLDPKV ARAKLFPGLSVHQALERVLRLPAVASKE LTNKVDRSVGGLVAQQCVCPLQTPLADL VVALSHEELIGAATALGEQFVKSLLDPKV ARAKLFGEGAALADACEAMVAVMALQV VVALSHEELIGAATALGEQFVKSLLDPKVCSGNWM WAAKLFGEGAALADACEAMVAVMALQV VOGGNSLSMARARVGTETVRAFGSLVISAY VCPDITATVTPDLKIPFGGGGHLLVYVALSFGG HRLGGTALAQCFSQLGEHPPDLYRCSGNWM WAAKLFGEGAALADACEAMVAVMALDPNLVS FSITQGLLKDRLLCSGRIUSDGGLVTCLLE AFAGNCGLVOVPVPRVDVSLTAEEPEGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGG AQPHAMVINSVNIGAVLEPVGGLARU TSFGJORLQAFBRCVAEEBKGRENGFSY LPPTFFKASVPREPGGSSRVALIREEGSNGEL EAADAFHLAGFEVWDVTMQLSGAGGLD FRGVAFVGGFSYADVLGSAKGWAAAVTTF RAGAELRFFKRRPDTTSLGVONGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRN SGRYESRWASVRVOPGPALMGGSCDG HAVMPHERAVRPWQAWRPPPFDTLTT PWLQLFINARNWTLEGGS 1345 2695 A 10396 65 642 GVGGFWAGTMASRAGFRAAGTDGSDFQHR RVAMHYQMSVTLKYELKKLIVYHLVIWLLL LAKMSVGHLRLLSHDQVAMPVGGIGSCDG HAVMPHERAVRPUGAGLARDALPN VGSMEMPFAAQQLVRHGKAYRFLFGFSAV MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK DDFVRCGRQSAAMGASARLLRAVINGAGG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSSQDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVSSSTTHFELKHLSGDLARDMARG GGGVAKAFIDGGKLDDDWMTRLALHELN TQYSMLLLGGFRTLIKLKUNYPL NYPEVLIGRLTARWHPASAGRT GGETNIKWPY NYPEVLIGRLTA	1		}]	<u> </u>]	
LALEKANQELGIALDSWDLDFYTKRÇQU HVT QKLVHSLJESIMSTQESSNPNNVLKFCDNS QKSLVHSLJESIMSTQESSNPNNVLKFCDNS QKSLVHSLJESIMSTQESSNPNNVLKFCDNS QKSEVFFLRFEDTRSRFQQQQGLRNVY AETHNFTGVCFFSGATTGTGGRIRDVQCTI RGAHVVAGTAGVCFGNLHIPGYNLPWEDI QYPGNFARPLEVALEASINGASDYGNKFGEP LAGFARSLGLQLPDGQRREWIKPIMFSGGI MEADHISKEAPFGMEVVKYGGPYYRIGVG GAASSVQVQGDNTSDLDFGAVQRGDFENW KMMRVIRACVEAPKGMPICSIMO GAASSVQVQGDNTSDLDFGAVQRGDFENW KMMRVIRACVEAPKGMPICSIMO GAASSVQVQGDNTSDLDFGAVQRGDFENW KMMRVIRACVEAPKGMPICSIMO GAASSVQVQGDNTSDLDFGAVQRGDFAN KMMRVIRACVEAPKGMPICSIMO GAPYQESNALLLRSPNRDFLTHVSARRGCDA PTPPTPVDLELEWVLGKMRRKEFFLQKKY MOPLALPPGLSWHQALEVLAPAVASKR LTNKVDRSVGGLVAQQCVCPLQTTLAALSIN QCPLALPPGLSWHQALEVLAPAVASKR LTNKVDRSVGGLVAQQCVCPLQTTLAALSIN VALSHEELIGAATALGEQPVFRNVGASKA LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR VALSHEELIGAATALGEQPVFRLAPAVASKR LTNKVDRSVGGLVAQQCVCPLQTTLAAVSKR VALSHEELIGAATALGEQPVFRLAPAVASKR VCPDITATVTDLKHPEGRGHLLYVALSPGC HALAQCFSQLGEHPPDLDLPRNLVR FSTTQGLLKDRLLCSGHDVSDGGLVTCLLEA AFAGNCGLQVDVPVPRVDLSVLFAEEPGG LEVQEPDLAQVLKRYRADAGLHCLEIGHT AFTGGLLKDRLLCSGHDVSDGGLVTCLLALW VCGDPNEDAAEMGPDSQPARGELLARIN TSFQLDRLQAEPRCVAEEERGLEERKGPSNGF LEVQEPDLAQVLKRYRADAGLHCLEIGHT ARGAELRFRKRPDTPSLGVCNGCQLLALLW WCGDPNEDAAEMGPDSQPARGELLARIN SGRYESRWASVRVOPGPALMGEGSNGF EMADAFHLAGFEVWDVTMQDLCSGAGGLED HAVMPHERAVRPWQWWWRPPFDTLTT PWLQLFINARNYTLESGS CHAVWFREPCOFSPRVAGTGCSCDLG HLAVMPHERAVRPWQWWWRPPFDTLTT PWLQLFINARNYTLESGS CHAVWFRLSSHDQVAMFUGGSCLG HLAVMPHERAVRPUGWSTYLLSHMSMGLISHJV YGSMEMPFAAQQLYRHCKAYRTLFGFSAV MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK LPSLLGLLSFFRNNISYLVLSMISMGLISHPTP YGGDDLTGEPLUQREDDVMTRLALHELKS GGGVVSSRTTHFELKHLSSGDLRDNMLRG GGGVVSSRTTHFELKHLSSGDLRDNMLRG GGGVVSSRTTHFELKHLSGDLRDNMTRLALHELKS GGGVVSSRTTHFELKHLSGDLRDNMTRLALHELKS GGGVVSSRTTHFELKHLSGDLRDDVMTRLALHELKS GGGVVSSRTTHFELKHLSGDLRDDVMTRLALHELKS GGGVVSSRTTHFELKHLSGDLRDDVMTRLALHELKS GGGVVSSRTTHFELKHLSGDRYNGETNKUWPY VGGDDLLGDPLUQREDDVRTVVKLKLKSVFT GGKKHK GGGCTVSSRTTHFELKHLSGDRYNGETNKUWPY UN			1		ļ		
NPSTVEAPDLAQSNSEHSRHWFFKGQLHVU QKLVHSILFESINTGESSIPPNIKFCDNSS IQGKEVRRLPPEDPTRSREQQQGLRHVVY AETHNPFIGVCFPSGATTGTGGRIRDVQCT RGAHVVAGTAGVCFGNLHIPGYNLFWEDLS QYPGNFARPLEVAIEASNGASDYGNKFGGP LAGGARSLGLQPDGQRREWIKPMFSGGIG MEADHISKEAPEPGMEVVKVGGPVYRIGE GAASSVQVGGDONTSDLPGAGVQGGCPGWE KMRVIRACVEAPKGNPICSLHDQGAGGNN NVLKELSDPAGAHIYTSRFQLGAGNIYTSRFQLGAGGCCGGAGLAGAGNITTSRFQLGAGNIYTSRFQLGAGCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1	j	ļ	1	ł	i	
QKL/HSL/FSSIMSTQESSPNNVLK/FCDNSS IQGKEVRI.PREPDTRESPROQQQGLRIVVY AETHNIPTGVCPSGATTGTGGRIRDVQCTT RGAHVVAGTAG YCFGNLHIPGYNLPWEDL: QYPGNRARPLEVAHEASNGASDYGNEFGEPP LAGFARS.GLQ.PDGQREWIKPIMPSGGIG MEADHISKEAPPEOMEVYGGPYNIGWG GAASSVQVQGDNTSDLDFGAVQRGGPYNIGWG GAASSVQVQGDNTSDLDFGAVQRGGPYNIGWG GAASSVQVQGDNTSDLDFGAVQRGGPYNIGWG GAASSVQVQGDNTSDLDFGAVQRGGPYNIGWG KMNRVIRACVAPKGNIPTSHVAARERGP, CFVGTITGDRRIVL.YDDRECPVRRNQGDA PTPPTTVDLLELEWVLGKMPKEFFLORFF MLQPLALPFGLSVHQALERVLR.PAVASKR LTINKVDRSVGGLVAQQQCVGFLQTPLADW. VVALSHEELIGAATALGGPYKSLLDFKVA. RLAVABALTINLVFALVTDLRDVKCSGNWM WAAKLPGGAALADACEAMYAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTDLKHPEGGAAMYAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTDLKHPEGGAAMAVAMALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTDLKHPEGGAAMAVAMALGV PSTITGGLKRALCSGHDVSDGGLVTCLLER AFAGNCGLQVDVPVRPVNVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGG AGPHAMVYSVNGAVVLEPVGELRALW TSSQLDRLQAEPRCVAEERGLRERMGPSY LPPTPKASVPREPGGPSPVLUREGSONGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADULGSAKGWAAAVTFH RAGAELRRFRKPPDTTSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLHNN SGRYESRWASVRVGPGAALMLRGEGSNGL PROVAFVGGFSYADULGSAKGWAAAVTFH RAGAELRRFRKPPDTTSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLHNN SGRYESRWASVRVGPGAALMLRGEMGAVL VWSAHGGGYVAFSSPELQAQIEARGLAPLH ADDDGNIFTEGYPLNNNEGSFORT RVAMHYQMSVTLKYEIKLIVYHLVIWLLL LIKLGLLSFPRNNISVLVLSMISMGLESIAPL YGSMEMFPAAQQLYRRIGKAYRFLFGFSAV MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWHTRAGGRVTNIEFPFPT TVGIDDLTGEFLIQREDDKPRTLAKLAYEI QTKPV-EVYQKKOVLETFSGTETNIKWPYV NVPFEVIKQRLTARWHTRAGGRVYNEFPPTT TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGIDDLTGEFLIQREDDKPRTLARLAFET TVGID		1	ļ	İ	ļ	!	
I I I I I I I I I I I I I I I I I I I			Į				
AETHNPTGVCPFSGATTGTGGRIRDVCTTC RGAHVVAGTAGVCFGNLHPGYNLPWEDLL QYPONFARPLEVAIEASNGASDVGNFGPP LAGFARSLGLQLPDGQRFEWIKPMFSGGG MEADHISKEAPPFOMEVYKVGGPVPTGVGV GAASSVQVQGDNTSDLDFGAVQRGDPME KMNRVIRACVAPKCMPICSLHDQGAGGNC NVLKELSDPAGAIIYTSRFQLGDPTLANAEIV GAEVQESNALLLRSPNRDFJTHVSARERCP- CFVGTITGDRRIVLVDDRECPVRRNGQGDA PTPPFTPVDLEEWVLGKMEFFELQRKP- MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQQCVGPITADV. VVALSHEELIGAATALGEGPVKSSLDPKVA- RLAVABALTNLVFALVTDLRDVKCSGNWM WAAKLPGGAALADACEAMVAVMAALCV VDGGKDSLSMARRVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGGRHLLVYALSPGG HRLGGTALAQCFSQLGEHPPDLDLPPNLVR. FSITQGLLKDRALLCSSGHDVGGLVTCLLER AFAGNCGLQVDVPPRVDVLSVLFAEEPCL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSY LPPTPFRASYPREPGGPSPVALREGSONGC EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAGLRRFRKRPDTTSLGVCNGCQLLALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGGALALL WVGGDPPEDAAEMCPPSTGGVAGGCCLALL WVGGDPPEDAAEMCPPSTGGTGATGCTGG HLAVMPHPERAVRPWQWWRPPPFDTT.TT PWLQLFINARWVTLEGSC HAVMPHPERAVRPWQWWRPPPFDTT.TT PWLQLFINARWVTLEGSC WGGWSGTTMSTTJVLIVSMIGLFSIPPI FGGLLSGFTAAGTDGSGFGRAGGTDGSFGGRAGGTGGGGGG HLAVMPHPERAVRPWQWWRPPFDTT.TT PWLQLFINARWVTLEGSC GYGGFWAGTTMSTRLILPLAVLIJWLLLL AKMSVGHLRLSHDQVAMPYQWSTPYLLS LPSLLGLLSFFRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAV MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK GGGYVSRTTTHEELKHL SGGDLLRDMLRGG EIGULAKAFIDQGKLIPDDVMTRLAHELKN TQYSWLLDGFPRTLPQAEDAGATGGDFGTFT TYGIDDLTGEFLIQREDDKPRTVRARLAKTE GIGVPS-WYCKKGVLTARWHPASGRVYNIEFNPPF TYGIDDLTGEFLIQREDDKPRTVRARLAKTE QTKPVLTYVQKKOVLGTFSGTETNKWPVVCW QTKPVLETYQKKKGVLTARWHPASGRVYNIEFNPPF TYGIDDLTGEFLIQREDDKPRTVTGRAGATGTFTGTETTNKWPTVCW QTKPVLETYQKKKGVLTARWHPASGRVYNIEFNPPF TYGIDDLTGEFLIQREDDKPRTVTGKARLAKT			l	1	1		QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
RGAHVVAGTAGYCFGNLHIPGYNLPWEDL QYPONPARPLEVAIEASNGSYGNKFGFP LAGFARSLGLQLPDGQRREWIKPIMFSGGIO MEADHISKEAPEPGMEVVKVGGPYYRIGVG GAASSVQVGQDNTSDLDFGAVQRGDFEME KMRAVIRACVEAPKGNPICSLHDQGAGGNO NVLKELSDPAGAIIYTSRFQLGDPTLNALEIY GAEYQESNALLLRSPNRDFLTHVSARBCP, CFVGTTTGDRRIVLYDDREPVPRRIGQGDA PTPPPTPVDLELEWVLGKMPRKEFFLQRKPI MLQPLALPPGLSVHQALERVLRLPAVASKR LTINKVDRSVGGLVAQQCHQTPLADV VVALSHEELIGAATALGEQPVKSLLDPKVA. RLAVAEALTAILVFALVTDLSVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTEVARGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPG HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKGDAVLCXSGNDVA VCPDITATVTPDLKHPEGRGHLLYVALSPG HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLKDQAVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVLEEPVGELRALWE TSFQLDRLQAPERCVAEEEGGLREEMGPSYV LPPTPPKASVPREPGGPSPRVALLREEGSNGE EMADAFHLAGFEV WDTVMQDLEELGHTGE AGPHAMVRVSVNGAVLEEPVGELRALWE TSFQLDRLQAPERCVAEEEGGLREEMGPSYV LPPTPPKASVPREPGGPSPRVALLREEGSNGE EMADAFHLAGFEV WDTVMQDLEAGLALL WVGGDPNEDQAAEMGPDSQARGGLALLAG WVGGDPNEDQAAEMGPDSQARGGLALLAG WVGGDPNEDQAAEMGPDSQARGGLALLAG WVGGDPNEDQAAEMGPDSQARGGLALLAG WVGGDPNEDQAAEMGPDSQARGGLALAG VWGGDPNEDQAAEMGPDSQARGGLALAGA VWGGDPNEDQAAEMGPDSQARGGLALAGA VWGGDPNEDQAAEMGPDSQARGGLAGLAD-H ADDDGNPTEQYPLNPNGSPGGVAGIGSCDG HLAVMPHPERAVRPWQWAWRPPFDTT.TT PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGFRAAGTDGSFOFIR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLIL AKMSVGHLRLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISLYLLSMISMGLFSIAP YGSMEMFPAAQQLYRHGKAYRFLFGFSAV MYLULVLAVQVHAWQLYYSKKLLDSWFTS GEKKHK GGGTVSSRITTHFELKHLSGDLLRDMALRG EIGVLAKAFIDQGKLPDDVMTRLALHELN TQYSWLLDGFPRTLPQAEATAGYDDTVM NVPFEVIKQRLTARWIHPASGRVYNIEFNPPF TVGIDDLTGEPLIQREDDKPBTVTRALHAFEN TVGSWLLDGFPRTLPQAEATAGYDDTVM NVPFEVIKQRLTARWIHPASGRVYNIEFNPPF TVGIDDLTGEPLIQREDDKPBTVTRALHAFEN TVGSWLLDGFPRTLPQAEATAGYDDTVM NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPBTTFGTETENNKWPYVD NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPBTTFGTETENNKWPYVD NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPBTTRALHAFEN TVGIDDLTGEPLIQREDDKPBTTRALHAFEN TVGIDDLTGEPLIQREDDKPBTTFGTETENNKWPYVD NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPBTTART	İ	i	ſ		ĺ		
QYPONFARPLEVAIEASNOASDYGNKFGEP LAGFARSLGIQLPDGQREWIKPINFSOGIO MEADHISKEAPEPGMEVVKVGGPVYRIGVC GAASSVQVQGOTSDLDFGAVQGDPEME KMNRVIRACVEAPKONPICSLHDQGAGGNC NVLKELSDPAGAIIYTSSFQLGDFTLNALEIY GAEVQESNALLLRSPNRDFLTHVSARERCP, CFVGTITGDRRIVLVDDRECPVRRNQQGDA PTPPTPYDLELEWVLGKMPRKEFTLQRKFY MLQPLALPPGLSVHQALERVLRLPAVASSR LTNKVDRSVGGLVAQQQCVGFQTPLADV. VVALSHEELIGAATALGCGPVLSLLDPKVA, RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACAMYAVMAALGV VDGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLVYALSPGG HRLGGTALAQCFSQLGGHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEB AFAGNCGQVDVPVPRVDVLSVLFAEEPGI LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSRQLDRLQABERCVAEERGLRERMGPSY LPPTPKASVPREPGGPSPRVAILREGSNOL EMADAFHLAGFEVWDTMODLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAVTHH RAGAELRERFKRPDTSTGJCVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SQRYESRWASVRVGPGFALMLRGMEGAVL VWSAHGEGVYAFSSFLQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLRHN SQRYESRWASVRVGPGFALMLRGMEGAVL VWSAHGEGVYAFSSFLQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGCSCDG HLAVMPHPERAVRPWQWAWPPPFDTLTT PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYYHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWSYYPLLS LPSLGGLLSFFRNNISYLVLSMISMGLESIAPL YGSMEMFPAAQQLYRHGKAYRRLFGFSAV MYLLVLAVQVHAWQLYYSKKLLDSWFTS LPSLGGLLSFFRNNISYLVLSMISMGLESIAPL YGSMEMFPAAQQLYRHGKAYRRLFGFSAV MYLLVLAVQVHAWQLYYSKLLLSWFTS QEKKHK QISVLAKARIDOGKLIPDDVMTRALHELKN TQYSWLLDGFPRTLAGAELDRAYQDTVIN NVPFEVIKQRLTARWHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETYTRILAYEIL QTRYNLEVLAQQCLTEFSGTETINKIWPYV OTRYPULLGFFRTNISYTRILAYEIT TVGIDDLTGEPLIQREDDKPETYTRILAYEIT TVGIDDLTGEPLIQREDDKPETYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQREDDKPETTYTRILAYEIT TVGIDDLTGEPLIQRED			ļ				`
LAGFARSLGLQLPDGQRREWIKPIMPSGGIG MEADHISKEAPEPGMEVVKVGGPVYRIGVV GAASSVQVQGDNTSDLDFGAVQRGDPEME KMNRVIRACVEAPKGNPICSJHDQGAGGNV NVLKELSDPAGAINTSQLGDFTLNALEIV GAEYQESNALLLRSPNRDELTHVSARERCP, CFYOTITIGPRIVLV LDBICEPVRRINGGDA PTPPPTPVDLELEWVLGKMPRKEFFLQRKFF MLQPLALPPGLSVHQALERVLRLPAVASKR LTINKVDRSVGGLVAQQCVGPLQPTLADV. VVALSHEELIGAATALGCQPVKSLLDPKVAA, RLAVBEALTNLVFALVILRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGG HRLGGTALAQCFSQLGGHPPDLDLPENLVR, FSITGGLLKDRLLSGFHDVSDGGLVTCLLEN AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLRRYRDAGLHCLEGHTGG AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEERGLRERMGPSYY LPPTPFKASVPREPGPSPVAILREGSONGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRERFKRRPTSTSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARRGLLRHN SGRYESREMSSVRVOPGPALMLRGMEGAVL WSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSTGPGAGICSCDG HLAVMPHFERAVRFWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLSHDQVAMPYQWEYPYLLS LPSLGLLSFFRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRRLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS LPSLGLLSFFRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRRLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRTTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDWMTRLALHELKN TQYSWLLUGGFRTLAPQAEALDRAYQDDTVIN NVPFEVIKQRLTARWHPASGRVYNEFPRPP TVGIDDLTGEPLIQREDDKPETVYNEFRPP TVGIDDLTGEPLIQREDDKPETVYNEFRPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPPP TVGIDDLTGEPLIQREDDKPETVYNEFRPP	l	- [1	[
MEADHISKEAPEPGMEVVKVGGPYPKIIGV GAASSVQVQGDNTSDLDFGAVQRGDPEME KMNRVIRACVEAPKGNPICSLHDQGAGGNC NVLKELSDPAGAIYTSRFQLGDFTLNALEIV GAEYQESNALLLRSPNRDFLTHVSARERCP, CFVGTITGBRRIVLVDDRECPVRRNQGGDA PTPPPTPVDLELEWVLGMPKREFFLQRKP MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQCVGHQTPLADV. VVALSHEELIGAATALGEQPVKSLLDPKVA, RLAVAEALTNLVFALVTDLRDVKSGNWM WAAKLPGEGAALADACAMVAWAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGGRILLLYVALSPG HRLGGTALAQCFSQLGEHPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEB. AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRSVNGAVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSY LPPTPKASVPREPGGPSPRVALREEGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAVTFH RAGAELRFKRRPDTSTSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARRGLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGVVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGIGSCDG HLAVMPHPERAVRRWQWAWPPPPDTLTT PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWSYPYLLS LPSLGGLLSFPRNNISYLVLSMISMGLESIAPL YGSMEMPFAAQQLYRHGKAYRRLFGFSAVI MYLVLVLAQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRNNMLRG EIGVLAKARIDOGGKLIPDDWTTRILAHELKN TQYSWLLJGPRTLAPQAEALDRAYQMDTVN NVPFEVIKQRLTARWHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETTYRRLAYEIL GYRPVLFVYQKKGVLETFSGTETNKIWPYV							QYPGNFARPLEVAIEASNGASDYGNKFGEPV
GAASSVQVQGDNTSDLDFGAVQRGDPEME KMNRVIRACVEAPKGNPICSLHDQGAGGNC NVLKELSDPAGAILYTSRFQLGDPTLNALEIV GAEYQESNALLLRSPNRDFLTHVSARERCP, CFVGTITGDRIVLYDDRECPVRRNGQGDA PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQCVGPLQTPLADV. VVALSHEELIGAATALGEQPVKSLLDPKVA. RLAVAEALTNLVATDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAWMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLVVALSPGG HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLKDRLLCSGHDVSDGGLVTCLLEN AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYPDAGLHCLELGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYY LPPTPFRASVPREPGGPSRVALIREGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADUCSAKGWAAAVTCH RAGAELRRFRKRPDTFSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESSRWASVRVGPGPALMLRGGGAVL WWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGMPTEQVPLNGSAKGWAAAVTCH ADDDGMPTEQVPLNFGSPGGVAGCCCGC HLAVMPHPERAVRPWQWAWRPPFDTILTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKLITYHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNINSYL VLSMISMGLFSLAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAV MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGYLAKAFIDQAKLIPDDVMTRLALHELKN TYSWLLDGFPRTLTPQAEALDRAYQDDTVIN NVPFEVIKQRL TARWHPASGRYVNIEFNPFP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYY					l .		LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS
KMMRVIRACVEAPKGNPICSLIHDQGAGGNC NVLKELSDPAGAIIYTSRFQLGDPTLNALEIU GAEYQESNALLIRSPNRDFLTHVSARERCP, CFVGTITGDRRIVLVDDRECPVRRNGQGDA PTPPPTPVDLELEWIGKMPRKEFFLORKP MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQCVOPLQTPLADV. VVALSHEELIGAATALGEQPVKSLLDPKVA. RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGALADACEMVAWAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLVVALSPGG HRLGGTALAQCESQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEB. AFAGNCGLQVDVPVPRVDVLSVLFAEEPGGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSGLDRLQAEPRCVAEEERGLRERMGPSY LPPTFPKASVPREPGGPSPRVALLREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGABELRRERKRPDTFSLGVONGCQLLALL WVGGDPNEDAAEMGPDSQPAPGGLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAGIEARGLAFLH ADDCOMPTEQYPLINPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWWAWPPPFDTTLTT PULQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLTYYHLVIWLLL LAKMSVGHRLLSHOVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYL VLSMISMGLESLAP, YGSMEMFPAAQQLYRHGKAYRFLFGPSAV MYLVLVAQVQHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMRLAHELKAYEI VGIDDLTGEPLIQREDDKPETVIKRLKAYEI VGIDDLTGEPLIQREDDKPETVIKRLKAYEI VGIDDLTGEPLIQREDKPETVIKRLKAYEI VGIDDLTGEPLIQR	}	l l	1	ļ	!]	MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
NVLKELSDPAGAIIYTSRFQLGDPTLNALEIY GAEYQESNALLLRSPINDFLTHVSARERCP, CFVGTTTGDRRIVLVDDRECPVRRNGQGDA PTPPPTPVDLELEWVLOKMPRKEFFLORKPP MLQPLAIPPGLSVHOLERVURLDPAVASKR LTNKVDRSVGGLVAQQCVGPLQTPLADV, VVALSHEELIGAATALGEQPVKSLLDPKVA. RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLVVALSPGG HRLGGTTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGF AGPHAMVRVSVNGAVVLEEPYGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSY LPPTPFKASVPREPGGPSRVALIREGSSNGE EMADAFHLAGFEVWDVTMQDLCSGGIGLD FRGVAFVGGFSY ADVLGSAKGWAAAVTFH RAGAELRRFRKPPDTSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPOPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAGIEAGLAPLH ADDGGMPTEQVPLNGSGGVAGGCSCDG HLAVMPHPERAVRPWQWAWRPPFFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLTYHLVIWLLL AKMSVGHZILSHAWGVHYPLLS UPSLLGLLSFPRNNISYLVLSMISMGLFSLAPL YGSMEMFPAAQQLYRHGKAYRFLGFSSA MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TYJSWLLDGFPRITLPQAEALDRAYQDDTVIN NVPFEVIKQRLTARWHPASGRYYNIEFNPPY TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYY			ļ	Ì			GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ
GAEVQESNALLLRSPNRDFLTHVSARERCP, CFVGTITGDRRIVLVDDRECPVRRNGQGDA PTPPTTVDLELEWVLGKMPRKEFFLQRKPF MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQQCVGPLQTPLADV, VVALSHEELIGAATALGEQPVKSLLDPKVA, RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGC HRLGGTALAQCFSQLGEHPPDLDLPENLVR, FSITQGLLKDRLLCSGHDVSDGGLVTCLLEN AFAGNCGLQVDYPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYY LPPTTPKASVPREPGGPSPRVAILREGSNGE EMADAFHL AGFEWDVTMGDLCSGAIGLD WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVQPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEGYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWWWRPPPFDTLTT PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYYHLVIWLLL AKMSYGHLRLISHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDITVIN NYPFEVIKQRLTARWHPASGRYYNIEPNPP TVGIDDLTGEPLIQREDDKPETVIKRKAYEI QTKRYULEYYQKEYTLKAYEI CHYPT.EVPLOTER CHYPTLE	1	1	l	1		1	KMNRVIRACVEAPKGNPICSLHDQGAGGNG
CFVGTTTGBRRIVLVDDRECPVRRNGQGDA PTPPPTPVDLELEWVLGKMPRKEFFLQRKPF MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQQCVGPLQTPLADV. VVALSHEELIGAATLGGEPVKSLLDPKVA. RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALDACCEMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGC HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEW AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSY LPPTPPKASVPREPGGPSPRVALIREGSNGG EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLRHN SGRYESRWASVRVGPGFALMLRGMEGAVL VWSAHGEGYVAPSSPELQAQIEARGLAPLH ADDDGNPTEQVPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGFRAAGTDGSDFOHR RVAMHYQMSVTKLEJEKLITVHLTVWLLJ AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVSMISMGLFSIAPL YGSMMFPAAQQLYRHGKAYRFLFGFSAVY YGSMMFPAAQQLYRHGKAYRFLFGFSAVY MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GRGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLTARWHPASGRYYNIEPNPPI TVGJDDLTGEPLIQREDDKPETVIKRLKAYEI QTKRYLLEGFEDKPETVIKRKYEI UTVGJDDLTGEPLIQREDDKPETVIKRLKAYEI GTKPVLEYYGLETTSGTETNKWPYV	l						NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW
PTPPPTPVDLELEWVLGKMPRKEFFLORKPF MLQPLALPPGLSVHQALERVLRLPAVASKR LTNKVDRSVGGLVAQQQCVGPLQTPLADV. VVALSHEELIGAATALGEQPVKSLLDPKVA. RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAWAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGC HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEN AFAGNCGLQVDVPRVDVULSVLFAEEPGGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPCVAEEERGLBERMGPSY LPPTPFKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPOSQARPGLLLRHN SGRYESRWASVROPDSQARPGLLLHHN ADDDGNPTEQYPLNPNOSPGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLISHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDIDTVIN NVPFEVIKQRLTARWIHPASGRYYNIEPNPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI GTKPVILEYPICKKGVLETTSGTETNKWPYV	l	- }	1				GAEYQESNALLLRSPNRDFLTHVSARERCPA
MLQPLALPPGLSVHQALERVLRIPAVÄSKR LTNKVDRSVGGLVAQQQCVGPLQTPLADV. VVALSHEELIGAATALGEGPVKSLLDPKVA. RLAVAEALTNILVFALVTDI.RDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGKDSLSMAARVGTETVRAPGSUSISY VCPDITATVTPDI.KHPEGRGHLLLYALSPGG HRLGGTALACPSQL GEHPPDLDI.PENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEN AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWF TSFQLDRLQAEPRCVAEERGLRERMGPSYV LPPTPKASVPREPGGFSPRVAILREGGSNGL EMADAFHLAGFEV WDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELREFRKRPDTFSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLRHN SGRYESRWASVRVOPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGIGSCDG HLAVMPHERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSYTLKYBIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGIFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEPNPPP TVGIDDLTGEPLIQREDDKPETVIKRIKAYEI GYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEPNPPP TVGIDDLTGEPLIQREDDKPETVIKRIKAYEI GYRFVLEYYQKKOVLETFSGTETNKIWPYV	1	- 1	ł	ĺ			CFVGTITGDRRIVLVDDRECPVRRNGQGDAP
LTNİKVDRSVGGLVAQQQCVGPLQTPLADV, VVALSHELIGAATALGEQPVKSLLDPKVA. RLAVAEALITULVFALVTDI.RDVKCSGNWM WAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGG HRLGGTALAQCFSQLGGHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEB. AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGG AGPHAMVRVSVNGAVVLEEPVGELRALWE TSSQLDRLQAEPRCVAEEERGLRERMGPSYY LPPTPPKASVPREPGGPSPRVAILREGGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTTSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQUEARGLAPLH ADDDGNPTEQPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVTWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSLAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEPNPPH TVGIDDLTGEPLIQREDDKPETVIKRILKAYPEI FVGIDDLTGEPLIQREDDKPETVIKRILKAYPEI FVGIDLTGEPLICKTER	ŀ	1					PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
VVALSHEELIGAATALĞEOPVKSLİDPKVA, RLAVABALTNLVFALVTDIRDVKCSGNWM WAAKLPGEGAALADACEAMYAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDILKHPGORGHLLYVALSPGG HRIGGTALAQCFSQLGEHPPDLDLPENLVR, FSITQGLLKDRLLCSGHDVSDGGLVTCLLEN AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWE TSSQLDRLQAEPRCVAEERGLRERMGPSYY LPPTPKASVPVGFOFSYNAULREGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFYGGFSYADVLGSAKGWAAAVTFH RAGAELRRERRDTTSSLGVCNGCOLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGFALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDGKGLIPDDWMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIENNPPT TVGIDDLTGEPLIQREDDKYPTVIKRIKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	ŀ			!			MLQPLALPPGLSVHQALERVLRLPAVASKRY
RLAVAEALTNI.VFALVTDI.RDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLL.YVALSPGG HRLGGTALAQCFSQLGEHPPDLDFENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRAL WE TSFQLDRL.QAEPRCVAEEERGLRERMGPSYY LPPTFPKASVPREPGGPSPRVALREEGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDOSDFOHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLISFPRNNISYL VLSMISMGLFSIAPI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRTTHHFELKHLSSGDLLRDNMLG EIGVLAKAFIDQGKLIPDDVMTRLAHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIERNPPY TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	ĺ	1	[Ì			LTNKVDRSVGGLVAQQQCVGPLQTPLADVA
WAAKLPGEGAALADACEAMVAVMAALGV VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATPDLKHPEGRGHLLYVALSPGG HRLGGTALAQCFSQLGEHPPDLDLPENLVR FSITQGLLKDRLLCSHDVSDGGLVTCLLEB AFAGNCGQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTFPKASVPREPGGPSFRVALLREEGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFKKRPDTFSLGVCNGCQLLALLW WVGGDPNEDAAEMGPDSOPARPGLLLRHN SGRYESRWASVRVOPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWWRPPFDTLTT: PWLQLFINANWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLESFRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPF TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	}	1					VVALSHEELIGAATALGEQPVKSLLDPKVAA
VDGGKDSLSMAARVGTETVRAPGSLVISAY VCPDITATVTPDLKHPEGRGHLLYVALSPGG HRLGGTALAQCFSOLGEHPPDLDLPENLVR. FSITQGLKDRLLCSGHDVSDGGLVTCLLER AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTPPKASVPREPGGPSPRVALIREEGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFFRRPDTTSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESSRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTIEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSYTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGFQSAAMGASRILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGYLLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRYYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	1	ı					RLAVAEALTNLVFALVTDLRDVKCSGNWM
VCPDITATVTPDLKHPEGRGHLLYVALSPGC HRLGGTALAQCFSQLGEHPPDLDLPERIVR. FSITGGLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGF AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYY LPPTFPKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESSRWSRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFOHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMPPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASRILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGYLLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEL QTKPVLEYYQKKGVLETFSGTETNKIWPYV		1		ŀ			WAAKLPGEGAALADACEAMVAVMAALGVA
HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEW AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTPFKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALLW WGGDPNEDAAEMGPDSQPARPGLLIRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQVPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPT TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI UTKPIDDLTGEPLIQREDDKPETVIKRLKAYEI UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET	ł						VDGGKDSLSMAARVGTETVRAPGSLVISAYA
HRLGGTALAQCFSQLGEHPPDLDLPENLVR. FSITQGLLKDRLLCSGHDVSDGGLVTCLLEW AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLEIGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTPFKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALLW WGGDPNEDAAEMGPDSQPARPGLLIRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQVPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPT TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI UTKPIDDLTGEPLIQREDDKPETVIKRLKAYEI UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGEPLIQREDTKPET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET UTKPINDLTGET	l		ļ				
FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSY LPPTPKASVPREPGGPSPRVAILREEGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPFFDTLTT PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPY TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV		-	Ì				
AFAGNCGLQVDVPVPRVDVLSVLFAEEPGL LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSSQLDRLQAEPRCVAEEERGLRERMGPSYY LPPTFPKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKPDTFSLGVCNGCQLLALLA WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEDAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH' ADDDGNPTEQYPLNPNGSPGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFOHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	ł	- 1	l	1			
LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTFPKASVPREPGGPSPRVAILREGSNGL EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLLALL WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNFTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYL VLSMISMGLFSIAPI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRTTTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWIHPASGRYYNIEFNPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	1	ŀ		<u>l</u>			
AGPHAMVRVSVNGAVVLEEPVGELRALWE TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTFPKASVPREPGGPSPRVAILREEGSNGD EMADAFHLAGFEV WDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLALLA WVGGDPNEDAAEMGPDSQPARPGLLLRIN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	}	ł		ł			,
TSFQLDRLQAEPRCVAEEERGLRERMGPSYG LPPTFPKASVPREPGGPSPRVAILREEGSNGE EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVLI VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLIRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV		- 1					1
LPPTFPKASVPREPGGPSPRVAILREEGSNGD EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLHT ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYL VLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	i	ļ					
EMADAFHLAGFEVWDVTMQDLCSGAIGLD FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYL VLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPP TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	}	ļ	1	}		j	
FRGVAFVGGFSYADVLGSAKGWAAAVTFH RAGAELRRFRKRPDTFSLGVCNGCQLLALLA WVGGDPNEDAAEMGPDSQPARPGLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARILRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	1						
RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGAPLHI ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYL VLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVLEYYQKKGVLETFSGTETNKIWPYV	1	Į.					
WVGGDPNEDAAEMGPDSQPARPGLLLRHN SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLH ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV		1					•
SGRYESRWASVRVGPGPALMLRGMEGAVL VWSAHGEGYVAFSSPELQAQIEARGLAPLHI ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLAHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	1	[,
VWSAHGEGYVAFSSPELQAQIEARGLAPLHI ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPFFDTLTT: PWLQLFINARNWTLEGSC GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARLIRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV				i i			•
ADDDGNPTEQYPLNPNGSPGGVAGICSCDG HLAVMPHPERAVRPWQWAWRPPPFDTLTT PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFFRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLAHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPF TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	Ì	- 1					
HLAVMPHPERAVRPWQWAWRPPPFDTLTT: PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLK YEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	ļ	1					
PWLQLFINARNWTLEGSC 1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV		- 1			i i		
1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHR RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV		- 1		ŀ			
RVAMHYQMSVTLKYEIKKLIYVHLVIWLLL AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	1345	2695	A	10396	65	642	
AKMSVGHLRLLSHDQVAMPYQWEYPYLLS LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	1	1 -0,5	''	.05,0		· · -	
LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPL YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV		1					
YGSMEMFPAAQQLYRHGKAYRFLFGFSAVS MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 I 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRA EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVILEYYQKKGVLETFSGTETNKIWPYV	l	1			1		I DOLL CLI CEDDYNLEAN AL SYLOWCI EGIYDI I
MYLVLVLAVQVHAWQLYYSKKLLDSWFTS QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	Ι.	·					
QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	ł						
1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	1				ļ		` `
GKGTVSSRITTHFELKHLSSGDLLRDNMLRG EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	1245	3607	 	10200	 	710	
EIGVLAKAFIDQGKLIPDDVMTRLALHELKN TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	1340	2090	^	10398	,	/10	
TQYSWLLDGFPRTLPQAEALDRAYQIDTVIN NVPFEVIKQRLTARWIHPASGRVYNIEFNPPI TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	(ĺ	[
NVPFEVIKQRLTARWIHPASGRVYNIEFNPPH TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	1	1					
TVGIDDLTGEPLIQREDDKPETVIKRLKAYEI QTKPVI.EYYQKKGVLETFSGTETNKIWPYV	l						
QTKPVI.EYYQKKGVLETFSGTETNKIWPYV					ļ		
					ļ		7
I I I I I I I I I I I I I I I I I I I	İ	1				ļ	•
	 , , ,	1000		10100			AFLQTKVPQRSQKASVTP
	1347	2697	Α	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL
VANPEALKILSAITQPVVVVAIVGLYRTGKS			لـــــا				VANPEALKILSAITQPVVVVAIVGLYRTGKSY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	•	!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]				peptide		/=possible nucleotide deletion, \=possible
1		ì	ĺ	sequence		nucleotide insertion
	.					LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
1		l '				PHPKKPEHTLVLLDTEGLGDVKKGDNONDS
i 1		Ì	[į .	WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY
						VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
l i						WTLRDFSLDLEADGOPLTPDEYLEYSLKLTO
						GTSOKDKNFNLPRLCIRKFFPKKKCFVFDLPI
1			}			HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
1						FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
] .]			GDLPCMENAVLALAQIENSAAVQKAIAHYD
		i i				QQMGQKVQLPAETLQELLDLHRVSEREATEV
						YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
						QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
í l						PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL OTYLKSKESVTDAILOTDOILTEKEKEIEVEC
						`
						VKAESAQASAKMVEEMQIKYQQMMEEKEKS
, ,			}			YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
1		ł i				QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI
1348	2698	À	10404	5	892	
1346	2096	A	10404	,	692	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
]						
						DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
						RFIERAPL VPK VRREPKNLSDIRGPSTEATEFT
Í !		ĺ				EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
	'	1				
1 1						GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
						FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
1 1						RERNNQNPWTFERIATANMLGIRKVLSPYDL
1240	2600		10100	50	1104	THKGKYWGKFYMPKRV
1349	2699	A	10409	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
						AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
1 1				;		NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
					İ	ALIDCLINPLQEQMEEWKKVANQLDKDHAK
)]						EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
						PQLDSALQDVNDKYLLLEETEKQAVRKALIE
ļ l						ERGRECTEISMLRPVIEEEISMLGEITHLQTISE
] .				'		DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
'						YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
						GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
1 1						DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
1350	0700		10410	F11	060	PDPNGGGPTTASGPPAAAEEAQRPRSM
1350	2700	A	10410	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
ļ [RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
j						AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
1 1						KYTYPALREEAPREHVESFFQKMDRNKDGV
1 [VTIEEFIESCQKDENIMRSMQLFDNVI

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

- 13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.